Anisotropic diffusion in image processing

Tomi Sariola

Advisor: Petri Ola

September 16, 2019

Abstract

Sometimes digital images may suffer from considerable noisiness. Of course, we would like to obtain the original noiseless image. However, this may not be even possible. In this thesis we utilize diffusion equations, particularly anisotropic diffusion, to reduce the noise level of the image. Applying these kinds of methods is a trade-off between retaining information and the noise level. Diffusion equations may reduce the noise level, but they also may blur the edges and thus information is lost.

We discuss the mathematics and theoretical results behind the diffusion equations. We start with continuous equations and build towards discrete equations as digital images are fully discrete. The main focus is on iterative method, that is, we diffuse the image step by step. As it occurs, we need certain assumptions for these equations to produce good results, one of which is a timestep restriction and the other is a correct choice of a diffusivity function.

We construct an anisotropic diffusion algorithm to denoise images and compare it to other diffusion equations. We discuss the edge-enhancing property, the noise removal properties and the convergence of the anisotropic diffusion. Results on test images show that the anisotropic diffusion is capable of reducing the noise level of the image while retaining the edges of image and as mentioned, anisotropic diffusion may even sharpen the edges of the image.

Contents

Introduction

Digital image consists of pixels. Each pixel contains a value given to describe its brightness or gray-level in grayscale images. Image size or resolution describes how many pixels there are in the image. Images with bigger resolution are sharper and closer to real world objects. Bits describe how many shades of gray images has. For example 8-bit image has $2^8 = 256$ different shades of gray. More bits mean that the difference between adjacent gray values is less noticeable. In this thesis images are 8-bit greyscale images.

Image noise is unwanted random variations in brightness level. In Gaussian noise these unwanted variations are normally distributed. There may also more specific noise types required by some applications. We use examples which are degraded intentionally with random noise.

Naturally, given the noisy image we want to obtain noiseless image or at least reduce the noise level. This thesis focuses on methods utilizing partial differential equation, more precisely diffusion type equations. These equations provide smoothed versions of initial image where noise has been reduced or removed. While reducing noise level it is desirable that smoothed image would contain features like edges and corners. Usually applying these equations means a trade-off between retaining information and noise level. If noise is reduced to minimum, finer details in image start to disappear. On the other hand, if these details are to be preserved, noise level may not be reduced enough.

Chapter 1 handles few different diffusion equations used in image processing. First, we consider denoising using the heat equation, also known as Gaussian blur. Applying Gaussian blur to noised image reduces the noise significantly. However, Gaussian blur uses constant diffusivity and thus it blurs edges while reducing noise.

By using diffusion equation which uses diffusivity which depends on image, we can reduce diffusion at the edges. Equation proposed by Perona and Malik uses diffusivity which gets smaller when edge is detected. It acts like heat equation on constant areas, but restricts diffusion across edges, and thus saves the edges while noise is reduced. Moreover, if diffusivity is chosen accordingly, an edge may even be enhanced. While

this equation is capable of reducing noise as well as heat equation, it does not perform well on noisy edges, that is, edge is preserved but so is noise.

Solution is to introduce anisotropic diffusion proposed by Weickert. Anisotropic diffusion reduces diffusion at the edges as does isotropic diffusion. Difference is that anisotropic diffusion still diffuses along edges and thus reduces noise at the edge.

Chapter 2 discusses theoretical results of diffusion equations in general. We introduce continuous diffusion equation, semi-discrete diffusion and discrete diffusion. We discuss well-posedness and other theoretical properties of these equations and also study under which assumptions these properties hold. In this chapter we represent construction of discrete diffusion equation for numerical approximation.

Chapter 3 addresses applications of diffusion filters. We first consider the design of diffusion filter and algorithm which is used for computation. The focus of thesis is to demonstrate what anisotropic diffusion is capable of so we will not cover performance of the algorithm. We will also not show any details of code as different programming languages have different methods and syntax.

Last section of chapter shows applications of diffusion equation. Images which are used consist of two very simple test images which are generated by computer, few photographs one of which is head shot and the other one is a tomography image. We consider edge-enhancing and noise removal properties of anisotropic diffusion and convergence as time grows. Anisotropic diffusion is also compared against Gaussian blur and isotropic diffusion. Lastly, we see that anisotropic diffusion can also be used as an artistic filter.

This thesis is inspired by the book *Anisotropic diffusion in image processing* authored by Weickert. It was originally doctoral thesis of Weickert. That book covers in much more detailed manner diffusion equations in image processing along with many other image restoration methods. Some of the proof are referred directly to his work. We encourage readers interested in fully detailed proofs to take a look of that book. Weickert also takes in consideration the efficient design of diffusion algorithm which is not covered in this thesis.

Chapter 1

Diffusion filtering

1.1 Physical background on diffusion

Diffusion is a physical process that equilibrates the density u of some quantity, such as heat, chemical concentration, pressure, etc. These physical phenomena have a mathematical representation in a form of diffusion equation, which we will derive from *Fick's first law of diffusion* and *the continuity equation*.

We denote partial derivative of u with respect to x_i by following notations:

$$\frac{\partial u}{\partial x_i} = \partial_{x_i} u = u_{x_i}. \tag{1.1}$$

Fick's first law of diffusion is the equation

$$j = -D \cdot \nabla u \tag{1.2}$$

It states that flux density j is proportional to concentration gradient ∇u defined as

$$\nabla u = (u_{x_1}, u_{x_2}) \in \mathbb{R}^2. \tag{1.3}$$

The negative sign arises because diffusion transfers substance opposite to the increasing concentration gradient. D is *diffusion tensor*, a positive definite matrix which describes relation between j and ∇u. If j and ∇u are parallel to each other, we may replace D by positive scalar. This case is called *isotropic*. In general j and ∇u are not parallel and this case is *anisotropic*. *Homogeneous diffusion* means that diffusion tensor is constant matrix and *inhomogeneous diffusion* means that diffusion tensor depends on coordinates. If diffusion tensor depends on evolving image itself we speak of *nonlinear diffusion* and if diffusion tensor does not depend on evolving image, it is called *linear diffusion*[1].

Continuity equation states that diffusion does not destroy or create substance and it is formulated by

$$u_t + \operatorname{div} j = 0, \tag{1.4}$$

where

$$\operatorname{div} j = \frac{\partial j_1}{\partial x_1} + \frac{\partial j_2}{\partial x_2}. \tag{1.5}$$

Combining Fick's law and continuity equation we obtain diffusion equation

$$u_t = \operatorname{div}(D \cdot \nabla u). \tag{1.6}$$

By replacing diffusion tensor with scalar 1 we obtain heat equation[2].

1.2 Diffusion filtering with heat equation

A simple way to smooth images is to let the initial image evolve via heat equation. This is known as a Gaussian blur. Consider the initial value problem

$$\begin{cases} u_t(x,t) = \Delta u(x,t) & \text{on} \quad \mathbb{R}^2 \times (0,T] \\ u(x,0) = u_0(x) & \text{on} \quad \mathbb{R}^2 \end{cases} \tag{1.7}$$

where

$$\Delta u(x,t) = \operatorname{div}\left(\nabla u(x,t)\right) = u_{x_1x_1} + u_{x_2x_2} \tag{1.8}$$

Above equation is called the heat equation. Notice that $u_0(x)$ is defined on whole $\mathbb{R}^2$. Initially u_0 is defined on bounded and closed rectangle $\Omega \subset \mathbb{R}^2$, that is, the image. However, we can extend u_0 to whole $\mathbb{R}^2$ by mirroring, that is, we use symmetry with respect to boundary of Ω. Note that mirroring preserves the continuity of u_0, but u_0 may not be differentiable after mirroring. We can now represent some results considering heat equation.

We define $u(x,t)$ by

$$u(x,t) = \frac{1}{4\pi t} \int_{\mathbb{R}^2} \exp\left(\frac{-|x-y|^2}{4t}\right) u_0(y)\, dy, \quad x \in \mathbb{R}^2, t > 0, \tag{1.9}$$

or equivalently as a convolution between Gaussian function with parameter $\sqrt{2t}$ and the initial value u_0

$$u(x,t) = \int_{\mathbb{R}^2} G_{\sqrt{2t}}(x-y) u_0(y)\, dy = \left(G_{\sqrt{2t}} * u_0\right)(x), \tag{1.10}$$

where

$$G_\sigma(x) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|x|^2}{2\sigma^2}\right), \tag{1.11}$$

is the Gaussian function or Gaussian kernel. Before introducing the theorem about solutions of heat equation we define sup-norm by the following:

$$||u(t)||_{L^\infty} = \sup_{x\in\mathbb{R}^2} |u(x,t)|. \tag{1.12}$$

The following theorem will consider the heat equation with a continuous initial image. We will discuss later discrete images domains, that is, digital images.

Theorem 1.1. *Let $u_0 \in C(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2)$ and u be defined by* (1.9)*. Then u is a smooth solution of* (1.7) *and*

$$\lim_{(x,t)\to(z,0)} u(x,t) = u_0(z), \tag{1.13}$$

for each $z \in \mathbb{R}^2$, $t > 0$. The solution u is unique, provided that u satisfies the growth estimate

$$|u(x,t)| \le Ae^{a|x|^2}, \quad x \in \mathbb{R}^2, \quad 0 \le t \le T, \tag{1.14}$$

for constants $a, A > 0$. Moreover, $u(x,t)$ satisfies the extremum principle

$$\inf_{x\in\mathbb{R}^2} u_0(x) \le u(x,t) \le \sup_{x\in\mathbb{R}^2} u_0(x). \tag{1.15}$$

and it depends continuously on the initial image.

Proof. We prove only the smoothness of the solution and continuous dependence on the initial image. We begin by proving the smoothness of (1.9). Partial derivatives of the $\exp(-|x|^2)$ are of the form

$$\partial_{x_1}^n \partial_{x_2}^m \exp(-|x|^2) = \sum_{\substack{k\le n\\ l\le m}} C_{k,l} x_1^k x_2^l \exp(-|x|^2) \tag{1.16}$$

for some $n, m \in \mathbb{N}$, from which we see that the Gaussian function is smooth.

Using symmetry we may rewrite

$$\int_{\mathbb{R}^2} \left|x_1^k x_2^l \exp(-|x|^2)\right| dx = 4\int_0^\infty \int_0^\infty x_1^k x_2^l \exp(-|x|^2)\, dx_1 dx_2. \tag{1.17}$$

Recall that the exponential grows faster than any polynomial which gives us that the integral above is finite. From this we deduce that all partial derivatives of the Gaussian are in $L^1(\mathbb{R}^2)$, that is

$$\partial_{x_1}^n \partial_{x_2}^m G_\sigma(x) \in L^1(\mathbb{R}^2) \tag{1.18}$$

for all $n, m \in \mathbb{N}$. This along with the dominated convergence theorem justifies us to take the derivatives inside integral and thus we get

$$\partial_{x_1}^n \partial_{x_2}^m u(x,t) = \int_{\mathbb{R}^2} \left(\partial_{x_1}^n \partial_{x_2}^m G_{\sqrt{2t}}(x-y)\right) u_0(y)\, dy, \tag{1.19}$$

so the solution of the heat equation is smooth. We use this formula to prove that equation (1.9) is really a solution of the heat equation. Differentiating under the integral sign we get

$$u_t(x,t) - \Delta u(x,t) = \int_{\mathbb{R}^2} \left[(\partial_t - \Delta) G_{\sqrt{2t}}(x-y)\right] u_0(y)\, dy, \quad \forall t > 0. \tag{1.20}$$

First, to evaluate the t-derivative one gets by straightforward differentiation

$$\frac{\partial}{\partial t}\left(\frac{1}{4\pi t}\exp\left(\frac{-|x-y|^2}{4t}\right)\right) = \exp\left(\frac{-|x-y|^2}{4t}\right)\cdot\left(\frac{-1}{4\pi t^2} + \frac{|x-y|^2}{16\pi t^3}\right). \tag{1.21}$$

The spatial derivatives of the Gaussian are

$$\begin{aligned} &\frac{\partial^2}{\partial x_i^2}\left(\frac{1}{4\pi t}\exp\left(\frac{-|x-y|^2}{4t}\right)\right) \\ =&\exp\left(\frac{-|x-y|^2}{4t}\right)\cdot\left(\frac{-1}{8\pi t^2} + \frac{(x_i-y_i)^2}{16\pi t^3}\right), \quad i \in \{1,2\}, \end{aligned} \tag{1.22}$$

and thus we obtain

$$\left(\frac{\partial}{\partial t} - \Delta\right)\left(\frac{1}{4\pi t}\exp\left(\frac{-|x-y|^2}{4t}\right)\right) = \left(\frac{\partial}{\partial t} - \Delta\right) G_{\sqrt{2t}}(x-y) = 0, \tag{1.23}$$

implying that the function u defined by (1.9) is truly a solution of the heat equation.

Next we prove the continuous dependence on the initial data. Suppose that $u(t)$ and $w(t)$ are solutions of the initial-value problem at $t \in (0,T]$ with u_0 and w_0 being the initial images respectively. Now

$$
\begin{aligned}
|u(x,t) - w(x,t)| &= \left| \int_{\mathbb{R}^2} G_{\sqrt{2t}}(x-y)(u_0(y) - w_0(y))\, dy \right| \\
&\leq \int_{\mathbb{R}^2} \left| G_{\sqrt{2t}}(x-y) \right| |u_0(y) - w_0(y)|\, dy \\
&\leq ||u_0 - w_0||_{L^\infty} \int_{\mathbb{R}^2} \left| G_{\sqrt{2t}}(x-y) \right| dy \\
&= ||u_0 - w_0||_{L^\infty}
\end{aligned}
\tag{1.24}
$$

for all $x \in \mathbb{R}^2$. In the calculations above we utilized the fact that

$$
\int_{\mathbb{R}^2} G_\sigma(x)\, dx = 1. \tag{1.25}
$$

Thus we get $||u(\cdot,t) - w(\cdot,t)||_{L^\infty} < ||u_0 - w_0||_{L^\infty}$ for all $t > 0$. Furthermore, any higher derivative of the solution depends continuously on the initial image. This follows from the (1.18).

The extremum principle follows from $u(x,t)$ satisfying the growth estimate

$$
|u(x,t)| \leq Ae^{a|x|^2}, \quad x \in \mathbb{R}^2, \quad 0 \leq t \leq T \tag{1.26}
$$

for constants $A, a > 0$. The uniqueness of the solution follows from the extremum principle. We refer to [2] for the fully detailed proof. □

To assume periodicity of the inital image is another way to guarantee uniqueness [4].

1.3 Non-linear models

Diffusion which depends on the evolving image itself is called non-linear diffusion. First, we study the Perona-Malik model without the regularization and then with the regularization.

1.3.1 Perona-Malik model

We consider filter proposed by Perona and Malik [3]. Because the heat equation blurs and mislocates edges, they proposed the following equation

$$
u_t = \operatorname{div}(g(|\nabla u|^2)\nabla u), \tag{1.27}
$$

where the decreasing function $g(s)$ should be chosen in such a way that $g(0) = 1$ and $g(s) \to 0$ as $s \to \infty$. They called this equation anisotropic diffusion filter. Note that this is inhomogeneous isotropic filter because of the scalar valued diffusivity.

With g chosen as proposed, model acts like the heat equation when ∇u is small. When ∇u is large, diffusion slows down, so $|\nabla u|^2$ acts as an edge detector. The idea of this model is to preserve edges of the image and this is achieved by diffusing less on the boundaries compared to the inner regions of the image.

Consider first the one-dimensional equation, and define $b(s) = 2sg'(s) + g(s)$. We write

$$\begin{aligned} u_t &= \frac{d}{dx}\left(g(u_x^2)u_x\right) \\ &= 2g'(u_x^2)u_x^2 u_{xx} + g(u_x^2)u_{xx} \\ &= b(u_x^2)u_{xx}. \end{aligned} \tag{1.28}$$

The function $g(s)$ can be chosen in such a way that, given a positive λ, the function $b(s)$ is positive for $s \leq \lambda$ and negative when $s > \lambda$. So when $s > \lambda$, equation acts like a backward heat-equation, meaning that the image is sharpening. However, as it turns out, we cannot say anything about the existence of the solution, let alone the stability. We will come back to this later.

Now suppose that u is smooth enough and write

$$\frac{\partial}{\partial t}(u_x) = \frac{\partial}{\partial x}(u_t) = \frac{\partial}{\partial x}\left(u_{xx}b(u_x^2)\right) = 2b'(u_x^2)u_x u_{xx}^2 + b(u_x^2)u_{xxx}. \tag{1.29}$$

Consider the point (x_0, t) at some positive time t, where u_x has its maximum. At this point $u_{xx} = 0$ and $u_{xxx} < 0$. If we have defined $g(s)$ so that $b(s) < 0$ when $s > \lambda$ then

$$\frac{\partial}{\partial t}(u_x) > 0, \tag{1.30}$$

provided that $u_x^2 > \lambda$. In this case, the edge is enhancing.

Let us now investigate the two-dimensional model. We start by noting that

$$\begin{aligned} \operatorname{div}(g(|\nabla u|^2)\nabla u) &= \frac{\partial}{\partial x_1}(g(|\nabla u|^2)u_{x_1}) + \frac{\partial}{\partial x_2}(g(|\nabla u|^2)u_{x_2}) \\ &= 2g'(|\nabla u|^2)\left(u_{x_1}^2 u_{x_1x_1} + 2u_{x_1}u_{x_2}u_{x_1x_2} + u_{x_2}^2 u_{x_2x_2}\right) \\ &+ g(|\nabla u|^2)\Delta u. \end{aligned} \tag{1.31}$$

Recall that

$$2g'(s) = \frac{b(s) - g(s)}{s}. \tag{1.32}$$

We use this to calculate

$$\begin{aligned}\operatorname{div}(g(|\nabla u|^2)\nabla u) =& \frac{b(|\nabla u|^2) - g(|\nabla u|^2)}{|\nabla u|^2}\left(u_{x_1}^2 u_{x_1x_1} + 2u_{x_1}u_{x_2}u_{x_1x_2} + u_{x_2}^2 u_{x_2x_2}\right) \\ &+ \frac{g(|\nabla u|^2)}{|\nabla u|^2}\left(u_{x_1}^2 u_{x_1x_1} + u_{x_1x_1}u_{x_2}^2 + u_{x_2x_2}u_{x_1}^2 + u_{x_2}^2 u_{x_2x_2}\right) \\ =& \frac{g(|\nabla u|^2)}{|\nabla u|^2}\left(u_{x_1x_1}u_{x_2}^2 - 2u_{x_1}u_{x_2}u_{x_1x_2} + u_{x_2x_2}u_{x_1}^2\right) \\ &+ \frac{b(|\nabla u|)^2}{|\nabla u|^2}\left(u_{x_1x_1}u_{x_1}^2 + 2u_{x_1}u_{x_2}u_{x_1x_2} + u_{x_2x_2}u_{x_2}^2\right)\end{aligned} \tag{1.33}$$

Let now (η, ξ) be an orthonormal frame such that η is parallel to ∇u. Then for the second order directional derivatives one has

$$u_{\eta\eta} = \frac{u_{x_1x_1}u_{x_1}^2 + 2u_{x_1}u_{x_2}u_{x_1x_2} + u_{x_2x_2}u_{x_2}^2}{|\nabla u|^2} \tag{1.34}$$

and

$$u_{\xi\xi} = \frac{u_{x_1x_1}u_{x_2}^2 - 2u_{x_1}u_{x_2}u_{x_1x_2} + u_{x_2x_2}u_{x_1}^2}{|\nabla u|^2}. \tag{1.35}$$

Hence we can write

$$\operatorname{div}(g(|\nabla u|^2)\nabla u) = g(|\nabla u|^2)u_{\xi\xi} + b(|\nabla u|^2)u_{\eta\eta}, \tag{1.36}$$

which means that we can interpret the Perona-Malik model as a sum of diffusions in the direction of gradient and perpendicular to the gradient. Now let's see what kind of a function is suitable as a diffusivity function g. We want

$$\lim_{s\to\infty}\frac{b(s)}{g(s)} = 0 \iff \lim_{s\to\infty}\frac{sg'(s)}{g(s)} = -\frac{1}{2}, \tag{1.37}$$

which means that when the gradient is large, we diffuse more in the perpendicular direction. From the condition above we get that $g(s) \approx 1/\sqrt{s}$ as $s \to \infty$ [4]. One example of such a function proposed by Perona and Malik in their paper [3] is

$$g(s) = \frac{1}{1 + (s/\lambda)^2}, \quad \lambda > 0. \tag{1.38}$$

However, the well-posedness of the Perona-Malik problem is unclear. There exists a theorem stating that if the initial image is not infinitely differentiable, then there is no weak solution. This theorem and more discussion of ill-posedness can be found in [4] and [1].

1.3.2 Regularized Perona-Malik model

In [5] Catté and others proposed the regularized Perona-Malik model. Instead of using ∇u for detecting edges, they use ∇u_σ where

$$u_\sigma = G_\sigma * u \tag{1.39}$$

is a blurred version of u. As previously, we extend u to whole $\mathbb{R}^2$. We introduce few notations required by the next theorem.

Let $\Omega \subset \mathbb{R}^2$ be an open domain, and let $H^1(\Omega)$ be the usual first-order Sobolev space consisting of all $u \in L^2(\Omega)$ having their weak gradient in $L^2(\Omega)$. The corresponding norm is

$$||u||_{H^1(\Omega)} = \left(\sum_{|\alpha| \leq 1} \int_\Omega |D^\alpha u|^2 \right)^{\frac{1}{2}}. \tag{1.40}$$

The space $L^2([0,T], H^1(\Omega))$ is the space of functions $u : [0,T] \to H^1(\Omega)$ which are strongly measurable on $[0,T]$, such that for almost every $t \in [0,T]$ it holds that $u \in H^1(\Omega)$. The definition of strong measurability can be found at the appendix section of [2]. The norm for this space is defined as

$$||u||_{L^2(0,T,H^1(\Omega))} = \left(\int_0^T |u|^2_{H^1(\Omega)} \, dt \right)^{\frac{1}{2}}. \tag{1.41}$$

In similar manner $C([0,T], L^2(\Omega))$ is the space of continuous functions $u : [0,T] \to L^2(\Omega)$ equipped with the norm

$$||u||_{C([0,T],L^2(\Omega))} = \max_{[0,T]} ||u(t)||_{L^2(\Omega)}. \tag{1.42}$$

We also denote regular inner product in $\mathbb{R}^2$ by $\langle \cdot, \cdot \rangle$. With these notations we state the following theorem:

Theorem 1.2. *Let $\Omega = (0,1) \times (0,1)$. Suppose $g : \mathbb{R}_+ \to \mathbb{R}_+$ be a decreasing function with $g(0) = 1$ and $g(s) \to 0$ as $s \to \infty$. In addition, assume mapping $s \to g(\sqrt{s})$ is smooth. Let $u_0 \in L^2(\Omega)$. Then we have a unique function*

$$u(t,x) \in C([0,T], L^2(\Omega)) \cap L^2([0,T], H^1(\Omega)) \tag{1.43}$$

which solves

$$\begin{cases} u_t = \operatorname{div}(g(|\nabla u_\sigma|)\nabla u) & \text{on} \quad (0,T) \times \Omega \\ \langle \nabla u, n \rangle = 0 & \text{on} \quad (0,T) \times \partial\Omega \\ u(0,x) = u_0(x) & \text{on} \quad \Omega \end{cases} \tag{1.44}$$

in the weak sense, where n denotes the unit outer normal to $(0,T) \times \partial\Omega$. Moreover, $u \in C^\infty((0,T) \times \overline{\Omega})$.

For a detailed proof we refer to [5]. Compared to the original Perona-Malik diffusion the equation above now satisfies well-posedness. Also, a noisy image may have false edges created by the noise. As we use ∇u_σ for finding the edges these false detections are reduced. This gives us better noise reduction properties.

1.3.3 Anisotropic models

Next we introduce anisotropic diffusion equations. As we have seen, isotropic diffusion uses a scalar valued diffusivity and this means the diffusion happens in the direction of the gradient. We may replace the scalar-valued diffusivity function with a diffusion tensor $D \in \mathbb{R}^{2\times 2}$ to obtain equation

$$u_t = \operatorname{div}(D \cdot \nabla u), \tag{1.45}$$

where the matrix D depends on the edge-estimator ∇u_σ and can be construed to fulfill desired properties, such as, preserving edges while reducing noise. We cover two types of diffusion tensors in next section, of which the first will be discussed more in this thesis.

Edge-enhancing diffusion (Anisotropic regularization of the Perona-Malik process)

As in the regularized Perona-Malik model, we use

$$\nabla u_\sigma = \nabla(K_\sigma * u) \tag{1.46}$$

for detecting edges. If we would use a scalar valued diffusivity $g(s)$ where g is chosen as proposed earlier, we would expect very little diffusion at edges. This means that noisiness would remain at the edges, and thus it would be desirable to diffuse perpendicular to the gradient at the edge.

We construct a diffusion tensor with the eigenvectors v_1 and v_2 such that

$$v_1 \parallel \nabla u_\sigma \quad \text{and} \quad v_2 \perp \nabla u_\sigma, \tag{1.47}$$

with the corresponding eigenvalues λ_1 and λ_2 chosen by

$$\begin{aligned} \lambda_1(\nabla u_\sigma) &= g(|\nabla u_\sigma|^2) \\ \lambda_2(\nabla u_\sigma) &= 1, \end{aligned} \tag{1.48}$$

where the function $g(s)$ is chosen in such a way that $b(s) = 2sg'(s) + g(s) < 0$ when $s > \lambda$ for some threshold $\lambda > 0$. Thus the diffusion tensor D depends on ∇u_σ and we may write $D(\nabla u_\sigma)$. We see that as ∇u_σ becomes larger, diffusion becomes more and

more perpendicular to ∇u_σ. If we have that $\sigma > 0$, the eigenvector $v_1 = \nabla u_\sigma$ does not coincide with ∇u. If we let $\sigma \to 0$, by the definition of eigenvalues we obtain the original Perona-Malik model as

$$D \cdot \nabla u = \lambda_1 \nabla u. \tag{1.49}$$

So as long as $\sigma > 0$ the model is an anisotropic model. We discuss the diffusion equation as above and its properties in the last chapter of this thesis.

Coherence-enhancing diffusion

Along with the edge-enhancing diffusion, Weickert has introduced an anisotropic diffusion equation for smoothing one-dimensional objects.

We use the same eigenvectors as above. For enhancing line-like structures, smoothing should be in the direction of v_2. One example of choosing the eigenvalues of the diffusion tensor D is

$$\begin{aligned} \lambda_1(\nabla u_\sigma) &= 0 \\ \lambda_2(\nabla u_\sigma) &= \frac{\eta \left|\nabla u_\sigma\right|^2}{1 + \left(\left|\nabla u_\sigma\right| / \sigma\right)^2} \quad (\eta > 0). \end{aligned} \tag{1.50}$$

We see that diffusion constructed as above diffuses only perpendicular to ∇u_σ.

Chapter 2

Theoretical results on the diffusion equation

2.1 Continuous diffusion

This chapter discusses theoretical results of both continuous and discrete diffusion. The basic structure of the diffusion equation is

$$\begin{cases} u_t = \operatorname{div}\left(D(\nabla u_\sigma) \cdot \nabla u\right) & \text{on} \quad \Omega \times (0,T] \\ u = u_0 & \text{on} \quad \Omega \times \{t=0\} \end{cases} \tag{2.1}$$

where $\Omega := (0, a_1) \times (0, a_2)$ is a rectangular image domain. For the following theoretical results, we need some additional assumptions which will be represented in this chapter along with the corresponding results.

2.1.1 Theoretical results

We define a more general edge descriptor than the magnitude of the gradient. As we have discussed earlier we would prefer not to diffuse across edges. We define vector $d(\theta) = (\cos\theta, \sin\theta)$ and function $F(\theta) = (d(\theta) \cdot \nabla u)^2$ which is a dot product of $d(\theta)$ and ∇u. By the definition of dot product we get that $F(\theta)$ can be minimized when $d(\theta) \perp \nabla u$ and maximized when $d(\theta) \parallel \nabla u$. We rewrite $F(\theta)$ as

$$\begin{aligned} F &= d^T \nabla u \nabla u^T d \\ &= d^T \begin{bmatrix} u_{x_1}^2 & u_{x_1} u_{x_2} \\ u_{x_1} u_{x_2} & u_{x_2}^2 \end{bmatrix} d. \end{aligned} \tag{2.2}$$

We may use the matrix

$$\nabla u \nabla u^T = \begin{bmatrix} u_{x_1}^2 & u_{x_1}u_{x_2} \\ u_{x_1}u_{x_2} & u_{x_2}^2 \end{bmatrix} \tag{2.3}$$

to describe the orientations of the image. As in the regularized models, blurring the image with a Gaussian kernel G_σ gives

$$u_\sigma = K_\sigma * u. \tag{2.4}$$

Now our edge descriptor becomes

$$\nabla u_\sigma \nabla {u_\sigma}^T = \begin{bmatrix} u_{\sigma_{x_1}}^2 & u_{\sigma_{x_1}}u_{\sigma_{x_2}} \\ u_{\sigma_{x_1}}u_{\sigma_{x_2}} & u_{\sigma_{x_2}}^2 \end{bmatrix}. \tag{2.5}$$

The matrix above has orthogonal eigenvectors $v_1 \parallel \nabla u_\sigma$ and $v_2 \perp \nabla u_\sigma$ with the corresponding eigenvalues $|\nabla u_\sigma|^2$ and 0.

For averaging local information of the image Weickert has convolved the matrix $\nabla u_\sigma \nabla {u_\sigma}^T$ with a Gaussian kernel G_ρ componentwise to obtain the matrix

$$J_\rho(\nabla u_\sigma) = G_\rho * \left(\nabla u_\sigma \nabla {u_\sigma}^T\right). \tag{2.6}$$

This matrix may be used in coherence-enchancing diffusion. The theoretical results below utilizes the matrix above although in the later sections of this thesis we use only the matrix $\nabla u_\sigma \nabla {u_\sigma}^T$ because we focus on the edge preserving diffusion and thus we do not need the local information, that is, the corners, the lines etc.

Now that we have covered enough notations we may state the following theorem which can be seen as an anisotropic extension of the regularized Perona-Malik model.

Theorem 2.1. *[1] Let $\Omega = (0, a_1) \times (0, a_2)$ be a rectangular image domain. Assume that $u_0 \in L^\infty(\Omega)$, $\rho \geq 0$ and $\sigma, T > 0$. Consider the problem*

$$\begin{cases} u_t = \operatorname{div}(D(J_\rho(\nabla u_\sigma))\nabla u) & \text{on} \quad \Omega \times (0, T], \\ u(0) = u_0 & \text{on} \quad \Omega, \\ \langle D(J_\rho(\nabla u_\sigma))\nabla u, n\rangle = 0 & \text{on} \quad \partial\Omega \times (0, T] \end{cases} \tag{2.7}$$

where the diffusion tensor satisfies following conditions:

- $D \in C^\infty(\mathbb{R}^{2\times 2}, \mathbb{R}^{2\times 2})$
- $d_{12}(J) = d_{21}(J)$ *for all symmetric matrices* $J \in \mathbb{R}^{2\times 2}$
- *For all* $w \in L^\infty(\Omega, \mathbb{R}^2)$ *with* $|w(x)| \leq K$ *on* Ω*, there exists a positive lower bound* $\nu(K)$ *for the eigenvalues of* $D(J_\rho(w))$*.*

Then there exists a unique solution $u(x, t)$ in the weak sense satisfying

$$\begin{aligned} u &\in C([0, T], L^2(\Omega)) \cap L^2([0, T], H^1(\Omega)) \\ u_t &\in L^2([0, T], H^1(\Omega)) \end{aligned} \tag{2.8}$$

Moreover, $u \in C^\infty(\bar{\Omega} \times (0, T])$. Finally, the solution u depends continuously on the initial value u_0 with respect to the L^2-norm and it fulfills the extremum principle

$$\operatorname*{ess\,inf}_\Omega u_0 \leq u(t) \leq \operatorname*{ess\,sup}_\Omega u_0, \quad t \in (0, T]. \tag{2.9}$$

Proof.

1. **Existence, uniqueness and regularity**
   The work of Weickert provides a fully detailed proof, see for example [1]. The proof is an anisotropic extension of the isotropic case which is found in [5].

2. **Extremum principle**
   For this proof, we recall the integration by parts formula.

   $$\int_\Omega u_{x_i} v \, dx = -\int_\Omega u v_{x_i} \, dx + \int_{\partial\Omega} u v n_i \, dS. \tag{2.10}$$

   Let u be a solution of the continuous diffusion equation. Denote

   $$b := \operatorname*{ess\,sup}_\Omega u_0. \tag{2.11}$$

Let $G \in C^1(\mathbb{R})$ be a function with $G(s) = 0$ when $s \in (-\infty, 0]$ and $0 < G'(s) \leq C$ when $s > 0$. Define

$$H(s) := \int_0^s G(y)\, dy \quad y \in \mathbb{R}, \tag{2.12}$$

$$\varphi(t) := \int_\Omega H(u(x,t) - b)\, dx. \tag{2.13}$$

The Cauchy-Schwarz inequality gives us

$$\begin{aligned} \int_\Omega |G(u(x,t) - b)u_t(x,t)|\, dx &\leq \int_\Omega C\, |(u(x,t) - b)u_t(x,t)|\, dx \\ &\leq C\, ||u(t) - b||_{L^2(\Omega)}\, ||u_t(t)||_{L^2(\Omega)}\,. \end{aligned} \tag{2.14}$$

This estimate and the dominated convergence theorem justify the following differentiation:

$$\begin{aligned} \varphi_t &= \frac{d}{dt} \int_\Omega H(u(x,t) - b)\, dx \\ &= \int_\Omega \frac{d}{dt} H(u(x,t) - b)\, dx \\ &= \int_\Omega G(u(x,t) - b)u_t(x,t)\, dx \\ &= \int_\Omega G(u(x,t) - b)\, \mathrm{div}(D(J_\rho(\nabla u_\sigma(x,t)))\nabla u(x,t))\, dx \\ &= \int_{\partial\Omega} G(u(x,t) - b)\langle D(J_\rho(\nabla u_\sigma(x,t)))\nabla u(x,t), n\rangle\, dx \\ &- \int_\Omega G'(u(x,t) - b)\langle D(J_\rho(\nabla u_\sigma(x,t)))\nabla u(x,t), \nabla u(x,t)\rangle\, dx. \end{aligned} \tag{2.15}$$

Recall that inner product is always non-negative. In addition $G' > 0$ and the integral over $\partial\Omega$ is zero because of the Neumann boundary condition. Thus

$$\varphi'(t) \leq 0. \tag{2.16}$$

Note that

$$0 \leq \varphi(t) = \int_\Omega H(u(x,t) - b)\, dx \leq \int_\Omega H(u(x,t) - u_0(x))\, dx. \tag{2.17}$$

We have that

$$H(s) = \int_0^s G(y)\, dy \leq \int_0^s Cy\, dy = \frac{C}{2} s^2, \tag{2.18}$$

so

$$0 \leq \varphi(t) \leq \int_\Omega H(u(x,t) - f(x))\, dx \leq \frac{C}{2} \left|\left|u(t) - u_0\right|\right|^2_{L^2(\Omega)}. \tag{2.19}$$

Because $u \in C([0,T], L^2(\Omega))$, we get that $\left|\left|u(t) - u_0\right|\right|^2_{L^2(\Omega)} \to 0$ as $t \to 0$, from which we obtain $\varphi(0) = 0$. Now we have that $\varphi \in C(0,T)$, $\varphi(0) = 0$, $\varphi \geq 0$ and $\varphi'(t) \leq 0$ and from these observations it follows that $\varphi = 0$ when $t \in [0,T]$. Now

$$\varphi(t) = \int_\Omega H(u(x,t) - b)\, dx = 0. \tag{2.20}$$

This is true when

$$H(u(x,t) - b) = 0 \tag{2.21}$$

almost everywhere. This gives us

$$\int_0^{u(x,t)-b} G(y)\, dy = 0. \tag{2.22}$$

The above integral holds when $u(x,t) - b \leq 0$ almost everywhere. Because u is smooth, we have

$$u(x,t) - b \leq 0 \quad \text{on} \quad \bar{\Omega} \times (0,T], \tag{2.23}$$

which is the maximum principle. By applying the maximum principle with the initial image $-u_0$ we obtain the minimum principle.

3. **Continuous dependence on the initial image** Let u be a solution with the initial image $u_0 \in L^\infty$ and w a solution with the initial image $w_0 \in L^\infty$. We have the following estimate from [1] and [5]:

$$\frac{d}{dt}\left(\left|\left|u(t) - w(t)\right|\right|^2_{L^2(\Omega)}\right) \leq C \left|\left|\nabla u(t)\right|\right|^2_{L^2(\Omega)} \left|\left|u(t) - w(t)\right|\right|^2_{L^2(\Omega)} \tag{2.24}$$

for some constant $C > 0$. By the Gronwall's inequality

$$\left|\left|u(t) - w(t)\right|\right|^2_{L^2(\Omega)} \leq \left|\left|u_0 - w_0\right|\right|^2_{L^2(\Omega)} \exp\left(C \int_0^t \left|\left|\nabla u(s)\right|\right|^2_{L^2(\Omega)}\, ds\right). \tag{2.25}$$

We know that u is bounded by the extremum principle. The derivatives of u_σ may be written as

$$\partial_{x_i} u_\sigma = \partial_{x_i}(u * G_\sigma) = u * (\partial_{x_i} G_\sigma), \quad i = 1, 2. \tag{2.26}$$

Recall that G_σ and its derivatives belong to $L^1(\Omega)$ and thus by the Cauchy-Schwarz inequality ∇u_σ is bounded. Note that the bound depends on σ.

For the matrix D there exists a lower bound $\nu > 0$ for the eigenvalues. We can estimate as follows:

$$\begin{aligned}
&\int_0^t ||\nabla u(s)||^2_{L^2(\Omega)}\, ds \\
\leq &\int_0^T ||\nabla u(s)||^2_{L^2(\Omega)}\, ds \\
= &\int_0^T \int_\Omega \langle \nabla u(x,s), \nabla u(x,s)\rangle \, dx\, ds \\
\leq &\frac{1}{\nu}\int_0^T \left| \int_\Omega \langle D(J_\rho(\nabla u_\sigma))\nabla u(x,s), \nabla u(x,s)\rangle \, dx \right| ds \\
= &\frac{1}{\nu}\int_0^T \left| \int_\Omega u(x,s)\cdot \operatorname{div}(D(J_\rho(\nabla u_\sigma(x,s)))\nabla u(x,s))dx \right| ds \\
\leq &\frac{1}{\nu}\int_0^T \int_\Omega |u(x,s)|\,|u_t(x,s)|\, dx\, ds \\
\leq &\frac{1}{\nu}\int_0^T ||u(s)||_{L^2(\Omega)}\,||u_t(s)||_{L^2(\Omega)}\, ds \\
\leq &\frac{1}{\nu}\,||u||_{L^2(0,T,L^2(\Omega))}\,||u_t||_{L^2(0,T,L^2(\Omega))} \\
\leq &\frac{1}{\nu}\,||u||_{L^2(0,T,H^1(\Omega))}\,||u_t||_{L^2(0,T,H^1(\Omega))}\,.
\end{aligned} \tag{2.27}$$

Let $\epsilon > 0$ and choose

$$\delta = \epsilon \cdot \exp\left(\frac{-C}{\nu}\,||u||_{L^2(0,T,H^1(\Omega))}\,||u_t||_{L^2(0,T,H^1(\Omega))}\right). \tag{2.28}$$

If $||u_0 - w_0||_{L^2(\Omega)} < \delta$, then

$$||u(t) - w(t)||_{L^2(\Omega)} < \epsilon \quad \forall t \in [0,T]. \tag{2.29}$$

□

2.1.2 Properties of the solution

This section covers the properties of the solution of a continuous diffusion equation. We introduce invariance properties, non-enhancement of extremas and convergence as $t \to \infty$. For more detailed explanations see [1].

Let $u(t)$ be a solution to problem mentioned in theorem (2.1) at positive time t. We define an operator T_t by

$$T_t(u_0) := u(t), \quad t \geq 0. \tag{2.30}$$

The operator T_t has the following properties listed below:

- *Grey level shift invariance:*

$$T_t(0) = 0 \tag{2.31}$$
$$T_t(u_0 + C) = T_t(u_0) + C \quad \forall t \geq 0. \tag{2.32}$$

- *Reverse contrast invariance:*

$$T_t(-u_0) = -T_t(u_0) \quad t \geq 0. \tag{2.33}$$

- *Translation invariance:*
  Define translation τ_h by $(\tau_h f)(x) = f(x + h)$. The we have

$$T_t(\tau_h u_0) = \tau_h(T_t u_0). \tag{2.34}$$

- *Isometry invariance:*
  Let $R \in \mathbb{R}^{2\times 2}$ be an orthogonal transformation and define $Rf(x) := f(Rx)$. Then

$$T_t(Ru_0) = R(T_t u_0). \tag{2.35}$$

The fundamental idea of diffusion is that substance is neither created or destroyed. We show that this property holds by proving that the average gray-level does not change when anisotropic diffusion is applied.

Theorem 2.2. *Let Ω be an open and bounded image domain. Define the average grey-level μ of the initial image u_0 by*

$$\mu := \frac{1}{|\Omega|} \int_\Omega u_0(x)\, dx. \tag{2.36}$$

Let $u(x,t)$ be a solution of diffusion equation. Then

$$\frac{1}{|\Omega|} \int_\Omega u(x,t)\, dx = \mu, \quad t > 0, \quad |\Omega| > 0. \tag{2.37}$$

Proof. Define

$$I(t) = \int_\Omega u(x,t) dx \tag{2.38}$$

Now by the Cauchy-Schwarz inequality

$$|I(t) - I(0)| = \left| \int_\Omega u(x,t) - u_0(x)\, dx \right| \leq ||u(t) - u_0||_{L^2(\Omega)} |\Omega|^{1/2}. \tag{2.39}$$

Thus $I(t)$ is continuous at 0. Recall that $u_t \in L^2([0,T], H^1(\Omega))$. Now with the dominated converge theorem and Ω being a Lipschitz-domain, using the integration by parts one can calculate the derivative of $I(t)$ as follows:

$$\begin{aligned}\frac{dI(t)}{dt} &= \frac{d}{dt}\int_\Omega u(x,t)\,dx = \int_\Omega u_t(x,t)\,dx \\ &= \int_\Omega \operatorname{div}(D(J_\rho(\nabla u_\sigma))\nabla u)\,dx = \int_{\partial\Omega} \langle D(J_\rho(\nabla u_\sigma))\nabla u, n\rangle\,dS = 0.\end{aligned} \tag{2.40}$$

Thus $I(t)$ is constant. □

The theorem below states that the local extremas of the image at any $t > 0$ are not enhanced. This means that the local maximum would get smaller and the local minimum bigger, and thus no new artifacts are created.

Theorem 2.3. *[1] Let u be a solution of the problem mentioned in the theorem* (2.1) *and consider some $t_0 > 0$. Suppose that $\xi \in \Omega$ is a local extremum of $u(\cdot, t_0)$ with a non-vanishing Hessian. Then*

$$u_t(\xi, t_0) < 0, \qquad \text{if } \xi \text{ is the local maximum,} \tag{2.41}$$

$$u_t(\xi, t_0) > 0, \qquad \text{if } \xi \text{ is the local minimum.} \tag{2.42}$$

Proof. To simplify notation we denote $D(J_\rho(\nabla u_\sigma)) =: D = (d_{ij})$. Then we have

$$\begin{aligned} u_t &= \operatorname{div}(D\nabla u) \\ &\partial_{x_1}(d_{11}u_x + d_{12}u_y) + \partial_{x_2}(d_{12}u_x + d_{22}u_y) \\ &\sum_{i=1}^{2}\sum_{j=1}^{2}(\partial_{x_j} d_{ij})u_{x_i} + \sum_{i=1}^{2}\sum_{j=1}^{2} d_{ij}u_{x_i x_j}. \end{aligned} \tag{2.43}$$

Suppose that ξ is the local maximum at time $\theta > 0$ from which we get $\nabla u(\xi, \theta) = 0$ and thus

$$u_t = \sum_{i=1}^{2}\sum_{j=1}^{2} d_{ij}u_{x_i x_j} \tag{2.44}$$

at (ξ, θ). The real symmetric matrix D can be decomposed as $D = S\Lambda S^T$, where $S \in \mathbb{R}^{2\times 2}$ is an orthogonal matrix and $\Lambda := \operatorname{diag}(\lambda_1, \lambda_2)$ with λ_1 and λ_2 being the positive eigenvalues of D.

Because we assumed that (ξ, θ) is the local maximum it follows that the Hessian matrix $H := H(u(\xi, \theta))$ is negative-definite and thus $B := S^T H S$ is also negative-definite. This gives us that $b_{ii} < 0$ for $i = 1, 2$.

The trace of a matrix is similarity-invariant and thus

$$\begin{aligned} u_t(\xi,\theta) &= \mathrm{Tr}(DH) \\ &= \mathrm{Tr}(S^T H S S^T D S) \\ &= \mathrm{Tr}(B\Lambda) \\ &= b_{11}\lambda_{11} + b_{22}\lambda_{22} \\ &< 0. \end{aligned} \tag{2.45}$$

For the local minimum the Hessian matrix is positive-definite, otherwise the proof is the same as above. □

Next we state the theorem which says that a filtered image convergences to a constant image. We do not cover the proof as it requires knowledge of the Lyapunov functionals. However, we refer to [1] for the full proof.

Theorem 2.4. *[1] Let u be the solution of the theorem* (2.1) *and μ defined as in* (2.37)*. Then*

$$||u(t) - \mu||_{L^p(\Omega)} \to 0 \quad \text{as } t \to \infty \text{ for } p \in [1, \infty). \tag{2.46}$$

2.2 Semidiscrete diffusion filtering

As digital images are not continuous, one would like to approximate continuous diffusion equation. First we consider spatial discretization of a continuous equation which is constructed by the method of finite differences. Later we will use the spatially discretized equation to obtain the fully discrete equation both in time and spatial domain.

Let us consider a continuous image domain $\Omega = (0, a_1) \times (0, a_2)$ with $N = n_1 \cdot n_2$ pixels. From this we get pixel width $h_1 = a_1/n_1$ and pixel height $h_2 = a_2/n_2$. Pixel $p(i, j)$ represents a point $(x_i, y_j) \in \Omega$ with $x_i = (i - 1/2)h_1$ and $y_j = (j_i - 1/2)h_2$, where $1 \leq i \leq n_1$, $1 \leq j \leq n_2$. The notation $u_{i,j}$ is the value of u at $p(i, j)$.

As the pixels in digital images are usually squares, this leads to $h_1 = h_2$, often even choice $h_1 = h_2 = 1$ is made. We represent images as vectors rather than matrices. In this work this is done columnwise, that is, image vector is obtained by stacking second column behind first column, third column behind second and so on. Vector representation can be done also row-wise.

We denote the set of pixel indices by $J = \{1, \ldots, N\}$. Usually in programming languages indexing starts at zero rather than one, so the set of pixel indices can also be noted as $\{0, \ldots, N - 1\}$. This is the case for example in Python language which is used in this thesis.

2.2.1 Theoretical results

Let $J = \{1, \ldots, N\}$. We consider a discrete image as a vector $u_0 \in \mathbb{R}^N$, $N \geq 2$, where u_{0j} represents the grey value of pixel for all $j \in J$.

The following theorem requires irreducible matrices so we start by introducing them. We say that a matrix A is called irreducible if for any $i, j \in J$ there exist $k_0, \ldots, k_r \in J$ with $k_0 = i$ and $k_r = j$ such that $a_{k_p k_{p+1}} \neq 0$ for $p \in \{0, \ldots, r-1\}$. With $J = \{1, 2, 3\}$, from the following matrices below the matrix A is irreducible and the matrix B is not irreducible:

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 5 \\ 4 & 0 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 5 & 6 \\ 4 & 2 & 7 \\ 0 & 0 & 3 \end{pmatrix}. \tag{2.47}$$

Let $k_0 = 3$ and $k_r = 1$. Our two possible choices are either $\{k_0, k_1\}$ with $k_0 = 3$, $k_1 = 1$ or $\{k_0, k_1, k_2\}$ with $k_0 = 3, k_1 = 2, k_2 = 1$. We see that $b_{31} = 0$ and thus the choice $\{3, 1\}$ is not possible. The choice $\{3, 2, 1\}$ is not possible either because $b_{32} = 0$ and $b_{21} = 4$ which implies that the matrix B is not irreducible.

Let us now prove the irreducibility of the matrix A. We denote the case where $k_0 = a$ and $k_r = b$ by (a, b). For the diagonal entries a_{ii}, $i \in \{1, 2, 3\}$ it holds that $a_{ii} \neq 0$, and thus the cases (i, i) for $i \in \{1, 2, 3\}$ are proved. With $a_{12} = 2$ and $a_{23} = 5$ the cases $(1, 2)$,$(1, 3)$ and $(2, 3)$ are proved. In addition, having $a_{31} = 4$ the cases $(3, 1)$,$(3, 2)$ and $(2, 1)$ are proved. Thus A is irreducible.

Theorem 2.5. *[1] Let $u_0 \in \mathbb{R}^N$ be an initial image. Consider the problem*

$$\begin{cases} \dfrac{du}{dt} = A(u)u, \\ u(0) = u_0, \end{cases} \tag{2.48}$$

where the matrix $A \in C(\mathbb{R}^N, \mathbb{R}^{N \times N})$ has the following properties:

- *$A(B, \mathbb{R}^{N \times N})$ is Lipschitz-continuous for all $B \subset \mathbb{R}^N$, where B is bounded.*
- $a_{ji}(u) = a_{ij}(u) \quad \forall i, j \in J, \quad u \in \mathbb{R}^N$
- $\sum\limits_{j \in J} a_{ji}(u) = 0 \quad \forall i \in J, \quad \forall u \in \mathbb{R}^N,$
- $a_{ji}(u) \geq 0, \quad \forall i \neq j, \forall u \in \mathbb{R}^N,$
- *A is irreducible for all $u \in \mathbb{R}^N$.*

Then for every $T > 0$, the problem above has a unique solution

$$u(t) = C^1([0,T], \mathbb{R}^N). \tag{2.49}$$

The solution depends continuously on the initial image u_0 and the right-hand side of the ODE system, that is, if we have two problems with the right-hand sides and the initial values close enough to each other, then we know the solutions of these two problems are close to each other. In addition, the solution of the above problem satisfies the extremum principle:

$$\min_{j\in J} u_0 \leq u_i(t) \leq \max_{j\in J} u_0 \quad \forall i \in J, \quad \forall t \in [0,T]. \tag{2.50}$$

Proof.

1. **Local existence**

   Let $u \in C([0,T], \mathbb{R}^N)$ and let $\beta > 0$. Define $\phi(t,u) := A(u)u$. Thus $\phi(t,u)$ is continuous on

   $$B_0 := [0,T] \times \left\{ u \in \mathbb{R}^N \mid ||u||_\infty \leq ||u_0||_\infty + \beta \right\}. \tag{2.51}$$

   We see that B_0 is compact, thus there exists $c > 0$ such that

   $$||u||_\infty \leq c, \quad \forall (t,u) \in B_0. \tag{2.52}$$

   Let $t_0 = 0$ and define R_0 by

   $$R_0 := \left\{ (t,u) \mid t \in [t_0, t_0 + \min\left(\frac{\beta}{c}, T\right), ||u - u_0||_\infty \leq \beta] \right\} \subset B_0. \tag{2.53}$$

   Recall that $A(u)$ is Lipschitz-continuous for every bounded subset of $\mathbb{R}^N$. This means that $\phi(t,u)$ is Lipschitz-continuous on R_0. By the Picard-Lindelöf theorem [6], initial-value problem

   $$\begin{cases} \dfrac{du}{dt} = \phi(t,u) = A(u(t))u(t), \\ u(0) = u_0, \end{cases} \tag{2.54}$$

   has a unique solution with $t \in [0, \min\left(\frac{\beta}{c}, T\right)]$. Notice that this is a proof for the local existence. We will cover the proof of the global existence later.

2. **Extremum principle**

   For the full proof we refer to [1]. Idea of the proof is to assume that the semi-discrete problem has a unique solution on $[0, \theta]$. Result of the proof is that extremum principle holds for all $t \in [0, T]$ which can then be used to prove the global existence of the solution.

3. **Global existence**

   For this proof, we use the same notations and results as in the proof of local existence. We know that there exists a solution for semi-discrete problem with $t \in [0, \min(\frac{\beta}{c}, T)]$. Let $t_1 = t_0 + \min(\frac{\beta}{c}, T)$. Define $u_1 := u(t_1)$. Now we consider the following problem:

$$\begin{cases} \dfrac{du}{dt} = \phi(t, u) = A(u(t))u(t), \\ u(t_1) = u_1. \end{cases} \tag{2.55}$$

   Now $\phi(t, u)$ is continuous on

$$B_1 = [0, T] \times \left\{ u \in \mathbb{R}^N \mid ||u||_\infty \leq ||u_1||_\infty + \beta \right\}. \tag{2.56}$$

   By the extremum principle, $||u(t_1)||_\infty = ||u_1||_\infty \leq ||u_0||_\infty$ and thus $B_1 \subset B_0$. Moreover, $||\phi(t, u)||_\infty \leq c$ for all $(t, u) \in B_1$. Further, we can show that $\phi(t, u)$ is Lipschitz-continuous in

$$R_1 := \left\{ (t, u) \mid t \in [t_1, t_1 + \min(\frac{\beta}{c}, T), ||u - u_1||_\infty \leq \beta] \right\}. \tag{2.57}$$

   Again, by the means of Picard-Lindelöf theorem, we have a unique solution for the problem we consider with $t \in [t_1, t_1 + \min(\frac{\beta}{c}, T)]$. This means that we have a unique solution for $t \in [0, 2\min(\frac{\beta}{c}, T)]$, and the extremum principle holds in this interval too. As a result, we have proved the uniqueness and the extremum principle for the interval $[0, T]$.

4. **Continuous dependence**

   Suppose again that $u \in C'([0, T], \mathbb{R}^N)$ is the solution of the semi-discrete problem. We have that $A(u)u$ is Lipschitz-continuous for every bounded subset of $\mathbb{R}^N$. From the proof of the uniqueness we get that there exists α such that $\phi(t, u)$ is Lipschitz-continuous on

$$S_\alpha = \{(t, v) \mid t \in [0, T], ||u - v||_\infty \leq \alpha\}. \tag{2.58}$$

Then by the theorem of continuous dependence from [7] p. 145, for every $\epsilon > 0$ there exists $\delta > 0$ such that if $\tilde{\phi}(t,u)$ is continuous in S_α and the inequalities

$$\begin{gathered} \left|\left|\tilde{\phi}(t,v) - \phi(t,v)\right|\right|_\infty < \delta, \quad \text{for } ||u-v||_\infty \leq \alpha, \\ ||\tilde{u}_0 - u_0||_\infty < \delta \end{gathered} \tag{2.59}$$

are satisfied, the solution $\tilde{u}$ of the "perturbed" initial value problem

$$\begin{cases} \dfrac{d\tilde{u}}{dt} = \tilde{\phi}(t,\tilde{u}), \\ \tilde{u}(0) = \tilde{u}_0, \end{cases} \tag{2.60}$$

exists in $[0,T]$ and it satisfies the inequality

$$||\tilde{u} - u|| < \epsilon. \tag{2.61}$$

In other words, the solution u depends continuously on the initial image and the right-hand side $\phi(t,u)$.

□

Next we state theorems considering the average gray-level of filtered images and convergence to constant image as in the case of continuous diffusion. We begin with the conservation of the average gray-level.

Theorem 2.6. *Let*

$$\mu = \frac{1}{N}\sum_{j\in J} u_{0j} \tag{2.62}$$

be the average grey-level of the initial image u_0. Semi-discrete diffusion does not change the average grey-level, that is,

$$\frac{1}{N}\sum_{j\in J} u_j(t) = \mu, \quad \forall t \geq 0. \tag{2.63}$$

Proof. Let $u(t)$ be a solution of the semi-discrete problem. Recall that the matrix $A(u)$ is symmetric and its row sums are zero and thus also column sums are zero, that is,

$$\sum_{j\in J} a_{jk}(u) = 0 \quad k \in J. \tag{2.64}$$

Now

$$\frac{d}{dt}\sum_{j\in J} u_j = \sum_{j\in J}\frac{du_j}{dt} = \sum_{j\in J}\sum_{k\in J} a_{jk}(u)u_k = \sum_{k\in J} u_k \sum_{j\in J} a_{jk}(u) = 0. \tag{2.65}$$

Since the derivative of the sum of pixel gray-values with respect to the time-variable is zero, the gray-level does not alter after applying diffusion. □

Theorem 2.7. *[1] Let $u(t) \in \mathbb{R}^N$ be a solution of the semi-discrete problem and let $\boldsymbol{\mu} \in \mathbb{R}^N$ be an average-grey value vector of the initial image. Then*

$$\lim_{t\to\infty} u(t) = \boldsymbol{\mu}. \tag{2.66}$$

Proving the convergence to a constant image would require the knowledge of Lyapunov functions and thus we do not cover the proof here, however, the proof can be found in [1].

2.2.2 Discretization of continuous diffusion

Isotropic case

First we cover the discretization of the isotropic diffusion, namely the regularized Perona-Malik diffusion. Recall that this can be written as

$$\begin{aligned}\frac{\partial u}{\partial t} &= \operatorname{div}\left(g(|\nabla u_\sigma|^2)\nabla u\right)\\ &= \frac{\partial}{\partial x_1}\left(g(|\nabla u_\sigma|^2)u_{x_1}\right) + \frac{\partial}{\partial x_2}\left(g(|\nabla u_\sigma|^2)u_{x_2}\right).\end{aligned} \tag{2.67}$$

From now on, for simplicity we will denote $g := g(|\nabla u_\sigma|^2)$. The following approximations will be used for the spatial derivatives:

$$\delta^*_{x_1} u_{i,j} = \frac{u_{i+\frac{1}{2},j} - u_{i-\frac{1}{2},j}}{h} \quad \text{and} \quad \delta^*_{x_2} u_{i,j} = \frac{u_{i,j+\frac{1}{2}} - u_{i,j-\frac{1}{2}}}{h}, \tag{2.68}$$

where $u_{i\pm\frac{1}{2},j\pm\frac{1}{2}}$ denotes the value of u at $(i \pm \frac{1}{2}, j \pm \frac{1}{2})$, which is obtained by interpolation. This kind of discretization resembles central difference. As we will see, some of these terms will cancel out.

The isotropic diffusion equation (2.67) at pixel (i, j) discretisizes as

$$\begin{aligned}\frac{du_{i,j}}{dt} &= \delta^*_{x_1}(g_{i,j}\delta^*_{x_1}u_{i,j}) + \delta^*_{x_2}(g_{i,j}\delta^*_{x_2}u_{i,j})\\ &= \delta^*_{x_1}\left(g_{i,j}\frac{u_{i+\frac{1}{2},j} - u_{i-\frac{1}{2},j}}{h}\right) + \delta^*_{x_2}\left(g_{i,j}\frac{u_{i,j+\frac{1}{2}} - u_{i,j-\frac{1}{2}}}{h}\right)\\ &= \frac{g_{i+\frac{1}{2},j}u_{i+1,j} - g_{i+\frac{1}{2},j}u_{i,j} - g_{i-\frac{1}{2},j}u_{i,j} + g_{i-\frac{1}{2},j}u_{i-1,j}}{h^2}\\ &\quad + \frac{g_{i,j+\frac{1}{2}}u_{i,j+1} - g_{i,j+\frac{1}{2}}u_{i,j} - g_{i,j-\frac{1}{2}}u_{i,j} + g_{i,j-\frac{1}{2}}u_{i,j-1}}{h^2}\\ &= \frac{u_{i+1,j}g_{i+\frac{1}{2},j} + u_{i-1,j}g_{i-\frac{1}{2},j} + u_{i,j+1}g_{i,j+\frac{1}{2}} + u_{i,j-1}g_{i,j-\frac{1}{2}}}{h^2}\\ &\quad - u_{i,j}\left(\frac{g_{i+\frac{1}{2},j} + g_{i-\frac{1}{2},j} + g_{i,j+\frac{1}{2}} + g_{i,j-\frac{1}{2}}}{h^2}\right).\end{aligned} \tag{2.69}$$

By interpolation

$$g_{i\pm\frac{1}{2},j\pm\frac{1}{2}} = \frac{g_{i\pm1,j\pm1} + g_{i,j}}{2}, \tag{2.70}$$

which are substituted into the above equation to obtain

$$\begin{aligned}
\frac{du_{i,j}}{dt} &= \frac{1}{2h^2}[(g_{i+1,j} + g_{i,j})u_{i+1,j} + (g_{i-1,j} + g_{i,j})u_{i-1,j} \\
&\quad + (g_{i,j+1} + g_{i,j})u_{i,j+1} + (g_{i,j-1} + g_{i,j})u_{i,j-1} \\
&\quad - (g_{i+1,j} + g_{i-1,j} + g_{i,j+1} + g_{i,j-1} + 4g_{i,j})u_{i,j}] \\
&= \frac{1}{2h^2}[(g_{i+1,j} + g_{i,j})(u_{i+1,j} - u_{i,j}) \\
&\quad + (g_{i-1,j} + g_{i,j})(u_{i-1,j} - u_{i,j}) + (g_{i,j+1} + g_{i,j})(u_{i,j+1} - u_{i,j}) \\
&\quad + (g_{i,j-1} + g_{i,j})(u_{i,j-1} - u_{i,j})].
\end{aligned} \tag{2.71}$$

Now we will denote $k = p(i,j)$ for some arbitrary pixel. We denote the set of neighbor pixels of k along x_1- and x_2-axis by $N(k)$. We can now write the above equation in the form

$$\frac{du_k}{dt} = \sum_{l\in N(k)} \frac{1}{2h^2}(g_l + g_k)(u_l - u_k) \tag{2.72}$$

which can be also represented in a matrix form. We define matrix $A = (a_{kl})$ by

$$a_{kl} = \begin{cases} \dfrac{1}{2h^2}(g_l + g_k) & \text{if} \quad l \in N(k) \\ -\displaystyle\sum_{l\in N(k)} \frac{1}{2h^2}(g_l + g_k) & \text{if} \quad l = k \\ 0 & \text{else.} \end{cases} \tag{2.73}$$

Now we will verify that this matrix satisfies requirements of the semi-discrete diffusion.

Proof. Lipschitz-continuity follows from the facts that diffusivity function g and gradient magnitude $|\nabla u_\sigma|^2$ are smooth and every image $u \in \mathbb{R}^N$ being bounded. Assume that we have pixels k and l which are the neighbor pixels of each other. Then we have that

$$l \in N(k) \iff k \in N(l). \tag{2.74}$$

By this condition we have that the matrix A is symmetric. Vanishing row sums and non-negative off-diagonals are clearly satisfied by the definition of the matrix A. For the irreducibility proof we first take pixels $i, j \in J$. Assume that $i = j$. Now $a_{ii} \neq 0$. Thus we choose $k_0 = i = j = k_r$. Now assume that $i \neq j$. We may choose path $k_0, \ldots, k_r$ such that k_p and k_{p+1} are neighbors for $p = 0, \ldots, r-1$. Then $a_{k_p k_{p+1}} \neq 0$. □

The reason for using discretization as above is because of the requirement for non-negative off-diagonals. For example, using central-difference to approximate the derivatives, non-negative off-diagonals are not guaranteed. Recall that g is chosen such that $g(0) = 1$ and $g \to 0$ as $t \to \infty$, thus g is non-negative. Hence we can see that the above discretization satisfies non-negativity for off-diagonals.

Anisotropic case

Discretizing anisotropic diffusion is not as straightforward as discretizing isotropic diffusion. Problem arises from the fact that non-negative off-diagonals are not ensured. Next we introduce a theorem proposed by Weickert that ensures representation such that off-diagonals are non-negative.

Theorem 2.8. *[1] Let $D \in \mathbb{R}^{2\times 2}$ be a positive definite symmetric matrix with a spectral condition number*

$$\kappa := \frac{\lambda_1}{\lambda_2}, \tag{2.75}$$

where λ_1, λ_2 are the eigenvalues of D. Then there exists some $m \in \mathbb{N}$ depending on κ, such that $\operatorname{div}(D\nabla u)$ *has second-order non-negative discretization on $(2m+1) \times (2m+1)$ stencil.*

Proof. [1] We sketch the ideas of the proof. Consider some $m \in \mathbb{N}$ and a $(2m+1) \times (2m+1)$ stencil. The boundary pixels of this stencil define $4m$ principal (from center pixel to boundaries) orientations $\beta_i \in (-\frac{\pi}{2}, \frac{\pi}{2}]$, $i = -2m+1, \ldots, 2m$, where

$$\beta_i := \begin{cases} \arctan\left(\dfrac{i}{m}\right), & |i| \le m \\ \operatorname{arccot}\left(\dfrac{2m-i}{m}\right), & m < i \le 2m \\ \operatorname{arccot}\left(\dfrac{i-2m}{m}\right), & -2m+1 \le i < -m. \end{cases} \tag{2.76}$$

Let $J_m := \{1, \ldots, 2m-1\}$. We construct a partition of $(-\frac{\pi}{2}, \frac{\pi}{2}]$ into $4m-2$ subintervals $I_i, |i| \in J_m$ by

$$(-\frac{\pi}{2}, \frac{\pi}{2}] = \bigcup_{i=-2m+1}^{-1} \underbrace{(\theta_i, \theta_{i+1}]}_{I_i} \cup \bigcup_{i=1}^{2m-1} \underbrace{(\theta_{i-1}, \theta_i]}_{I_i}, \tag{2.77}$$

with

$$\theta_i := \begin{cases} 0, & i = 0 \\ \frac{1}{2}\arctan\left(\frac{2}{\cot\beta_i - \tan\beta_{i+1}}\right), & i \in \{1,\dots,2m-2\}, \beta_i + \beta_{i+1} < \pi/2, \\ \frac{\pi}{4}, & i \in \{1,\dots,2m-2\}, \beta_i + \beta_{i+1} = \pi/2, \\ \frac{\pi}{2} + \frac{1}{2}\arctan\left(\frac{2}{\cot\beta_i - \tan\beta_{i+1}}\right), & i \in \{1,\dots,2m-2\}, \beta_i + \beta_{i+1} > \pi/2, \\ \frac{\pi}{2}, & i = 2m-1 \end{cases} \tag{2.78}$$

and $\theta_i = -\theta_{-i}$. It can be verified that $\beta_i \in I_i$ for $|i| \in J_m$.

Now let $\lambda_1 \geq \lambda_2 > 0$ be the eigenvalues of D with the corresponding eigenvectors $(\cos\psi, \sin\psi)$ and $(-\sin\psi, \cos\psi)$, where $\psi \in (-\frac{\pi}{2}, \frac{\pi}{2}]$. Let $\psi \in I_k$ and $D = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Weickert proposes the following conditions:

- If the condition number is bounded by

$$\frac{\lambda_1}{\lambda_2} \leq \min(\cot(\rho_k - \beta_k)\tan\rho_k, \cot(\beta_k - \eta_k)\cot\eta_k) =: \kappa_{k,m}, \tag{2.79}$$

  where

$$\rho_k = \begin{cases} \theta_k, & |k| \in \{1,\dots,2m-2\}, \\ \frac{1}{2}(\theta_k + \beta_k), & |k| = 2m-1 \end{cases} \tag{2.80}$$

  and

$$\eta_k = \begin{cases} \frac{1}{2}\beta_k, & |k| = 1, \\ \theta_{k-1}, & |k| \in \{2,\dots,2m-1\}, \end{cases} \tag{2.81}$$

  then

$$\min(a - b\cot\beta_k, c - b\tan\beta_k) \geq 0 \tag{2.82}$$

- The upper bound for the condition number goes to infinity as m grows, that is,

$$\lim_{m\to\infty} \min_{|i|\in J_m} \kappa_{i,m} = \infty. \tag{2.83}$$

If these conditions hold, then a non-negative splitting of the diffusion equation is possible and we may rewrite the right-hand side of the diffusion equation as

$$\operatorname{div}(D\nabla u) = \partial_{\beta_0}(\alpha_0 \partial_{\beta_0} u) + \partial_{\beta_k}(\alpha_k \partial_{\beta_k} u) + \partial_{\beta_{2m}}(\alpha_{2m} \partial_{\beta_{2m}} u), \tag{2.84}$$

where ∂_{β_k} denotes a directional derivative in the direction of β_k and the non-negative coefficients α_0, α_k and α_{2m} are given by

$$\begin{aligned} \alpha_0 &= a - b\cot\beta_k, \\ \alpha_k &= \frac{b}{\sin\beta_k \cos\beta_k}, \\ \alpha_{2m} &= c - b\tan\beta_k. \end{aligned} \tag{2.85}$$

□

The above theorem states that if the stencil size is large enough, we can find a spatial discretization of the anisotropic diffusion equation which satisfies properties proposed earlier. Stencils used in image processing are usually of sizes 3×3 or 5×5 depending on the finite difference method used.

Next we show a spatial discretization with the stencil size being 3×3. As is in the proof we divide

$$(-\frac{\pi}{2}, \frac{\pi}{2}] = (-\frac{\pi}{2}, 0] \cup (0, \frac{\pi}{2}] =: I_{-1} \cup I_1. \tag{2.86}$$

We also obtain

$$\begin{aligned} \beta_1 &= \arctan(1) = \frac{\pi}{4} \\ \beta_{-1} &= \arctan(-1) = -\frac{\pi}{4}. \end{aligned} \tag{2.87}$$

Using (2.79) and symmetry, we obtain an upper bound for the condition number of D:

$$\frac{\lambda_1}{\lambda_2} \leq \kappa_{1,1} = \kappa_{-1,1} = 3 + 2\sqrt{2}. \tag{2.88}$$

Now let $\psi \in I_1$. Using (2.85), we get the non-negative diffusivities

$$\begin{aligned} \alpha_0 &= a - b\cot\beta_1 = a - b \\ \alpha_1 &= \frac{b}{\sin\beta_1 \cos\beta_1} = 2b \\ \alpha_2 &= c - b\tan\beta_1 = c - b. \end{aligned} \tag{2.89}$$

With a similar argument, for $\psi \in I_{-1}$ we get the diffusivities

$$\begin{aligned} \alpha_0 &= a - b\cot\beta_{-1} = a + b \\ \alpha_{-1} &= \frac{b}{\sin\beta_{-1} \cos\beta_{-1}} = -2b \\ \alpha_2 &= c - b\tan\beta_{-1} = c + b. \end{aligned} \tag{2.90}$$

We may combine the above diffusivities as

$$\begin{aligned} \alpha_{-1} &= |b| - b \\ \alpha_0 &= a - |b| \\ \alpha_1 &= |b| + b \\ \alpha_2 &= c - |b| \,. \end{aligned} \tag{2.91}$$

For the spatial discretization we write the divergence term as a sum of the directional derivatives:

$$\begin{aligned} \operatorname{div}(D\nabla u) &= \partial_{\beta_{-1}}(\alpha_{-1}\partial_{\beta_{-1}}u) + \partial_{\beta_0}(\alpha_0\partial_{\beta_0}u) + \partial_{\beta_1}(\alpha_1\partial_{\beta_1}u) + \partial_{\beta_2}(\alpha_2\partial_{\beta_2}u) \\ &=: H + I + J + K. \end{aligned} \tag{2.92}$$

We may discretize (2.92) in the same manner as we did isotropic diffusion equation (2.67). The discretization of H becomes:

$$\begin{aligned} H &\approx \delta^*_{\beta_{-1}}\left(\alpha_{-1_{i,j}}\delta^*_{\beta_{-1}}u_{i,j}\right) \\ &= \delta^*_{\beta_{-1}}\left(\alpha_{-1_{i,j}}\frac{u_{i+\frac{1}{2},j-\frac{1}{2}} - u_{i-\frac{1}{2},j+\frac{1}{2}}}{\sqrt{2}h}\right) \\ &= \frac{1}{\sqrt{2}h}\left(\frac{\alpha_{-1_{i+\frac{1}{2},j-\frac{1}{2}}}(u_{i+1,j-1} - u_{i,j}) - \alpha_{-1_{i-\frac{1}{2},j+\frac{1}{2}}}(u_{i,j} - u_{i-1,j+1})}{\sqrt{2}h}\right) \\ &= \frac{1}{2h^2}\alpha_{-1_{i+\frac{1}{2},j-\frac{1}{2}}}u_{i+1,j-1} + \frac{1}{2h^2}\alpha_{-1_{i-\frac{1}{2},j+\frac{1}{2}}}u_{i-1,j+1} \\ &\quad - \frac{1}{2h^2}\left(\alpha_{-1_{i+\frac{1}{2},j-\frac{1}{2}}} + \alpha_{-1_{i-\frac{1}{2},j+\frac{1}{2}}}\right)u_{i,j}. \end{aligned} \tag{2.93}$$

With the use of interpolation we obtain

$$\begin{aligned} H &\approx \frac{1}{2h^2}\frac{\alpha_{-1_{i+1,j-1}} + \alpha_{-1_{i,j}}}{2}u_{i+1,j-1} + \frac{1}{2h^2}\frac{\alpha_{-1_{i-1,j+1}} + \alpha_{-1_{i,j}}}{2}u_{i-1,j+1} \\ &\quad - \frac{1}{2h^2}\frac{\alpha_{-1_{i+1,j-1}} + \alpha_{-1_{i-1,j+1}} + 2\alpha_{-1_{i,j}}}{2}u_{i,j} \\ &= \frac{\alpha_{-1_{i+1,j-1}} + \alpha_{-1_{i,j}}}{4h^2}u_{i+1,j-1} + \frac{\alpha_{-1_{i-1,j+1}} + \alpha_{-1_{i,j}}}{4h^2}u_{i-1,j+1} \\ &\quad - \frac{\alpha_{-1_{i+1,j-1}} + \alpha_{-1_{i-1,j+1}} + 2\alpha_{-1_{i,j}}}{4h^2}u_{i,j}. \end{aligned} \tag{2.94}$$

The discretization of I, J and K is done in a very similar manner, only by changing the diffusivities and the directional derivatives to suitable ones. In the table below a full 3×3 -stencil is shown in order to approximate the diffusion equation at pixel

(i, j). The stencil represents the non-zero values of $A(u)$ at row $p_{i,j}$, with $p_{i,j}$ denoting the pixel index for pixel (i, j). The centermost entry of the table gives the value of $A(u)$ at index $(p_{i,j}, p_{i,j})$ and the upper left entry gives the value of $A(u)$ at index $(p_{i,j}, p_{i-1,j-1})$.

$\frac{\alpha_{-1_{i-1,j+1}} + \alpha_{-1_{i,j}}}{4h^2}$	$\frac{\alpha_{2_{i,j+1}} + \alpha_{2_{i,j}}}{2h^2}$	$\frac{\alpha_{1_{i+1,j+1}} + \alpha_{1_{i,j}}}{4h^2}$
$\frac{\alpha_{0_{i-1,j}} + \alpha_{0_{i,j}}}{2h^2}$	$-\frac{\alpha_{-1_{i-1,j+1}} + 2\alpha_{-1_{i,j}} + \alpha_{-1_{i+1,j-1}}}{4h^2}$ $-\frac{\alpha_{1_{i+1,j+1}} + 2\alpha_{1_{i,j}} + \alpha_{1_{i-1,j-1}}}{4h^2}$ $-\frac{\alpha_{2_{i,j+1}} + 2\alpha_{2_{i,j}} + \alpha_{2_{i,j-1}}}{2h^2}$ $-\frac{\alpha_{0_{i+1,j}} + 2\alpha_{0_{i,j}} + \alpha_{0_{i-1,j}}}{2h^2}$	$\frac{\alpha_{0_{i+1,j}} + \alpha_{0_{i,j}}}{2h^2}$
$\frac{\alpha_{1_{i-1,j-1}} + \alpha_{1_{i,j}}}{4h^2}$	$\frac{\alpha_{0_{i,j-1}} + \alpha_{0_{i,j}}}{2h^2}$	$\frac{\alpha_{-1_{i+1,j-1}} + \alpha_{-1_{i,j}}}{4h^2}$

Recall the principal directions

$$\beta_{-1} = -\frac{\pi}{4}, \quad \beta_0 = 0, \quad \beta_1 = \frac{\pi}{4}, \quad \beta_2 = \frac{\pi}{2} \tag{2.95}$$

and let $B = \{-1, 0, 1, 2\}$ be the set of principal direction indices of the stencil. Let $N_b(k)$ be the set of neighbors for pixel k in the direction β_b. Now the discretization of the anisotropic diffusion can be written in the same manner as isotropic discretization:

$$\frac{du_k}{dt} = \sum_{b \in B} \sum_{l \in N_b(k)} \frac{c_b}{2h^2} (\alpha_{b_l} + \alpha_{b_k})(u_l - u_k), \tag{2.96}$$

where

$$c_b = \begin{cases} 1, & \text{if} \quad b \in \{0, 2\}, \\ \frac{1}{2}, & \text{if} \quad b \in \{-1, 1\}. \end{cases} \tag{2.97}$$

By using the above discretization with a diffusion tensor that satisfies the requirement for a bounded condition number (2.79), we obtain a semi-discrete diffusion equation that satisfies the theoretical results proposed earlier in this chapter.

2.3 Discrete diffusion filtering

In the previous section we discussed a diffusion equation with a spatial discretization and with a continuous time variable. However, numerical approximations require also discretization in the time domain.

Theorem 2.9. *[1] Let $u_0 \in \mathbb{R}^N$ and $J = \{1, \ldots, N\}$. Consider the discrete problem*

$$\begin{cases} u^{(0)} = u_0, \\ u^{(k+1)} = Q(u^{(k)})u^{(k)}, \quad \forall k \in \mathbb{N}, \end{cases} \tag{2.98}$$

where the matrix Q has the following properties:

- $Q \in C(\mathbb{R}^N, \mathbb{R}^{N\times N})$,
- $q_{ij}(v) = q_{ji}(v), \quad \forall i, j \in J, \quad \forall v \in \mathbb{R}^N$,
- $\sum_{j\in J} q_{ij}(v) = 1, \quad \forall i \in J, \quad \forall v \in \mathbb{R}^N$,
- $q_{ij}(v) \geq 0, \quad \forall i, j \in J, \quad \forall v \in \mathbb{R}^N$,
- *The diagonal entries of Q are positive*
- *Q is irreducible.*

Then every $u_0 \in \mathbb{R}^N$ generates a unique sequence $\left(u^{(k)}\right)_{k\in\mathbb{N}}$. Also, for every finite k, $u^{(k)}$ depends continuously on u_0 and it satisfies the extremum principle:

$$\min_{j\in J} u_{0j} \leq u_i^{(k)} \leq \max_{j\in J} u_{0j}, \quad \forall k \in \mathbb{N}, \forall i \in J. \tag{2.99}$$

Proof. For this proof we assume that $u^{(k)}$ is the solution of the discrete problem.

1. **Uniqueness**
   Let $u^{(k)}$ and $v^{(k)}$ be solutions of the discrete diffusion problem with the same initial value. Define $w^{(k)} = u^{(k)} - v^{(k)}$. Then $w^{(0)} = 0$, from which follows that $w^{(k)} = 0$, that is, $u^{(k)} = v^{(k)}$.

2. **Continuous dependence**
   By the definition of continuity and from the fact that $Q(w)$, $w \in \mathbb{R}^N$ is continuous, it follows that $Q(w)w$ is continuous. Thus by induction, we obtain that
   $$u^{(k)} = Q(u^{(k-1)})u^{(k-1)} \tag{2.100}$$
   depends continuously on u_0.

3. **Extremum principle**
   Recall that the non-negative matrix Q has unit row sums. Let us consider some arbitrary pixel i. We have
   $$u_i^{(k+1)} = \sum_{j\in J} q_{ij}(u^{(k)})u_j^{(k)} \leq \max_{l\in J} u_l^{(k)} \sum_{j\in j} q_{ij}(u^{(k)}) = \max_{l\in J} u_l^{(k)}, \tag{2.101}$$
   and by iterating, we get
   $$u_i^{(k+1)} \leq \max_{l\in J} u_l^{(0)} = \max_{l\in J} u_{0l}. \tag{2.102}$$
   The minimum principle is proved as above, just by taking the minimum instead of the maximum and changing the direction of inequalities.

□

Next we state theorems for the fully discrete diffusion problem considering the conservation of the average grey-level and convergence to a constant image similar to the semi-discrete diffusion problem. Let us begin with the conservation of the average grey-level.

Theorem 2.10. *[1] Let*
$$\mu = \frac{1}{N}\sum_{j\in J} u_{0j} \tag{2.103}$$
be the average grey-level of the initial image. Then we have that
$$\frac{1}{N}\sum_{j\in J} u_j^{(k)} = \mu, \quad k \in \mathbb{N}_0 \tag{2.104}$$

Proof. By the symmetry of Q and its unit row sums we have
$$\begin{aligned}\sum_{j\in J} u_j^{(k)} &= \sum_{j\in J}\sum_{i\in J} q_{ij}(u^{(k-1)})u_i^{(k-1)} = \sum_{i\in J} u_i^{(k-1)} \sum_{j\in J} q_{ij}(u^{(k-1)}) \\ &= \sum_{i\in J} u_i^{(k-1)} = \cdots = \sum_{i\in J} u_i^{(0)} = \sum_{i\in J} u_{0i}.\end{aligned} \tag{2.105}$$
Thus discrete diffusion does not affect the average grey-level. □

Theorem 2.11. ***(Convergence to a constant image)****[1]*
Let $\boldsymbol{\mu} = (\mu, \ldots, \mu) \in \mathbb{R}^N$ *be the vector with values of the average grey value. Then*

$$\lim_{k\to\infty} u^{(k)} = \boldsymbol{\mu}, \quad \forall k \in \mathbb{N}. \tag{2.106}$$

Again, we do not prove the convergence to steady-state as it requires knowledge of the Lyapunov sequences, however, the full proof can be found in [1].

Suppose that we know a solution $u^{(k)}$ for the semi-discrete problem (2.5) at time $t = k\tau$, where $k \in \mathbb{N}$ and $\tau > 0$ is the size of the time-step. To approximate the left-hand side of the semi-discrete equation we use the following approximation:

$$\frac{du(t)}{dt} \approx \frac{u^{(k+1)} - u^{(k)}}{\tau}. \tag{2.107}$$

The following theorem shows that for a suitable time-step τ the above discretization leads to a scheme which satisfies the requirements as proposed in theorem (2.9).

Theorem 2.12. *[1]****(Semi-implicit schemes)***
Let $\alpha \in [0, 1]$, $\tau > 0$, *and let the matrix* $A(u^{(k)})$ *satisfy the requirements for semi-discrete diffusion presented in the previous section. Then the* α*-semi-implicit scheme*

$$\frac{u^{(k+1)} - u^{(k)}}{\tau} = A(u^{(k)}) \left(\alpha u^{(k+1)} + (1-\alpha)u^{(k)}\right) \tag{2.108}$$

fulfills the requirements of the discrete diffusion, if

$$\tau \leq \frac{1}{(1-\alpha)\max\limits_{i\in J} |a_{ii}(u^{(k)})|}, \quad \alpha \in (0, 1). \tag{2.109}$$

In explicit case when $\alpha = 0$, *the requirements are satisfied if*

$$\tau < \frac{1}{\max\limits_{i\in J} |a_{ii}(u^{(k)})|}, \tag{2.110}$$

and for the semi-implicit case $\alpha = 1$ *the requirements are satisfied for all* $\tau > 0$.

Proof. Because in this thesis we use an iterative scheme, we prove the explicit case $\alpha = 0$ only. The discrete diffusion equation then becomes

$$\frac{u^{(k+1)} - u^{(k)}}{\tau} = A(u^{(k)})u^{(k)} \tag{2.111}$$

from which we deduce

$$u^{(k+1)} = \left(I + \tau A(u^{(k)})\right) u^{(k)} =: B(u^{(k)})u^{(k)}, \tag{2.112}$$

where $I \in \mathbb{R}^{N\times N}$ is the identity matrix. Continuity, symmetry and non-negative off-diagonals follow easily from the properties of $A(u^{(k)})$. Because $A(u^{(k)})$ is Lipschitz-continuous it follows that $B(u^{(k)})$ is continuous. The symmetry of I and $A(u^{(k)})$ implies that $B(u^{(k)})$ is symmetric. The non-negative off-diagonals of $A(u^{(k)})$ imply that also $B(u^{(k)})$ has non-negative off-diagonals.

The row sums of $B(u^{(k)})$ are written as

$$\begin{aligned}\sum_{j\in J} b_{ij}(u^{(k)}) &= b_{ii}(u^{(k)}) + \sum_{j\in J, j\neq i} b_{ij}(u^{(k)}) \\ &= 1 + \tau a_{ii}(u^{(k)}) + \tau \sum_{j\in J, j\neq i} a_{ij}(u^{(k)}) \\ &= 1 + \tau \left(\sum_{j\in J} a_{ij}(u^{(k)}) \right).\end{aligned} \tag{2.113}$$

We recall that

$$\sum_{j\in J} a_{ij}(u^{(k)}) = 0, \quad \forall i \in J \tag{2.114}$$

and thus

$$\sum_{j\in J} b_{ij}(u^{(k)}) = 1, \quad \forall i \in J. \tag{2.115}$$

Because $A(u^{(k)})$ has non-negative off-diagonals, then

$$a_{ii}(u^{(k)}) = -\sum_{j\in J, j\neq i} a_{ij}(u^{(k)}) < 0. \tag{2.116}$$

The matrix $B(u^{(k)})$ has a positive diagonal if

$$1 + \tau a_{ii}(u^{(k)}) > 0, \quad \forall i \in J. \tag{2.117}$$

This holds when

$$\tau < \frac{1}{\max\limits_{i\in J} |a_{ii}(u^{(k)}|)}. \tag{2.118}$$

If the above time-step restriction holds, then the diagonal entries of $B(u^{(k)})$ are positive. Using similar reasoning as proving the irreducibility of $A(u^{(k)})$ we obtain the irreducibility of $B(u^{(k)})$. □

Chapter 3

Algorithm and examples

This chapter handles the applications of the diffusion equation. We discuss the design of the anisotropic diffusion, the algorithm and finally some examples.

3.1 Filter design

We want to design a filter which prevents diffusion across edges. Now let $v_1 \parallel \nabla u_\sigma$ be the eigenvector of D with the corresponding eigenvalue λ_1. We may choose

$$\begin{cases} \lambda_1 = g(\mu_1) \\ \lambda_2 = 1, \end{cases} \tag{3.1}$$

where $\mu_1 = |\nabla u_\sigma|^2$ and smooth function g is chosen in such a way that $g(0) = 1$ and $g(s) \to \infty$ as $s \to \infty$. Weickert uses a diffusivity function of type

$$g(s) := \begin{cases} 1, & s \leq 0 \\ 1 - \exp\left(\dfrac{-C_m}{(s/\lambda)^m}\right), & \text{else,} \end{cases} \tag{3.2}$$

where C_m is chosen in such a way, that flux function $b(s) = 2sg'(s) + g(s)$ is positive for $s \in [0, \lambda]$ and negative for $s > \lambda$ for chosen $m \in \mathbb{N}$. Perona and Malik proposed in their work [3] a diffusivity function of type

$$g(s) = \frac{1}{1 + (s/K)^2}, \tag{3.3}$$

where K is a constant to be chosen. Function

$$g(s) = \frac{1}{1 + \frac{1}{3}\left(\frac{s}{\lambda}\right)^2} \tag{3.4}$$

is used in this thesis inspired by the Perona-Malik diffusivity equation.

Recall the one-dimensional Perona-Malik equation

$$u_t = u_{xx} b(u_x^2), \tag{3.5}$$

where

$$b(s) = 2sg'(s) + g(s). \tag{3.6}$$

With $g(s)$ chosen as above we have that $b(s)$ is positive when $s \in [0, \lambda]$ and negative when $s > \lambda$. This means that edges are enhancing when $|\nabla u_\sigma|^2 > \lambda$.

Recall the results from the previous chapter that the theoretical properties such as extremum principle hold for discrete anisotropic diffusion if the matrix $A(u^{(k)})$ has positive diagonals. We saw that for 3×3 stencil this would mean that

$$\frac{\lambda_1}{\lambda_2} = \frac{1}{g(s)} \leq 3 + 2\sqrt{2} \iff g(s) \geq \frac{1}{3 + 2\sqrt{2}} \approx 0.1716. \tag{3.7}$$

However, with our choice of $g(s)$ this requirement is not satisfied and thus we cannot expect positive diagonals of the matrix $A(u^{(k)})$. Furthermore, the assumptions for the theoretical results do not necessarily hold. Later, as we will cover some examples we will see that despite the possible negative values in the diagonal anisotropic diffusion provides good visual results.

The reason to use the Perona-Malik diffusivity is because the Weickert type diffusivity function decreases too rapidly and thus the minimum of $b(s)$ becomes too small, that is, edge-enhancing is too radically. This lead to an unstable equation. On the other hand, Weickert uses semi-implicit schemes in his work. One hypothesis is that this may cause the better stability compared to the iterative scheme.

3.2 Algorithm

Recall the continuous anisotropic diffusion equation with the initial image u_0

$$\begin{cases} u_t = \operatorname{div}(D(J(\nabla u_\sigma))\nabla u) & \text{on} \quad \Omega \times (0, T], \\ u(x, 0) = u_0(x) & \text{on} \quad \Omega, \\ \langle D(J(\nabla u_\sigma))\nabla u, n\rangle = 0 & \text{on} \quad \partial\Omega \times (0, T], \end{cases} \tag{3.8}$$

which can be discretized as

$$\begin{cases} \dfrac{u^{(k+1)} - u^{(k)}}{\tau} = A(u^{(k)})u^{(k)} \\ u^{(0)} = u_0, \end{cases} \tag{3.9}$$

which in the iterative form is written as

$$\begin{cases} u^{(k+1)} = \left(I + \tau A(u^{(k)})\right) u^{(k)}, \\ u^{(0)} = u_0. \end{cases} \tag{3.10}$$

The matrix $A(u^{(k)})$ represent the discretization of the term $\operatorname{div}(D(J(\nabla u_\sigma))\nabla u)$ and it is construed by using the method of finite differences as discussed in the previous chapter. The blurred version u_σ of the image u is obtained by using the iterative scheme

$$u_{i,j}^{(k+1)} = \tau\left(u_{i+1,j}^{(k)} + u_{i-1,j}^{(k)} + u_{i,j+1}^{(k)} + u_{i,j-1}^{(k)}\right) + (1-4\tau)u_{i,j}^{(k)}, \quad \tau < 0.25 \tag{3.11}$$

which satisfies the requirements for discrete diffusion. The spatial derivatives of the image u are approximated by using the central differences

$$u_{x_{1i,j}} = \frac{u_{i+1,j} - u_{i-1,j}}{2} \quad \text{and} \quad u_{x_{2i,j}} = \frac{u_{i,j+1} - u_{i,j-1}}{2}. \tag{3.12}$$

As usual, the values of the pixels that do not fall into image domain are obtained by mirroring.

Let us now assume that we have chosen the diffusivity function $g(s)$ to be used along with the suitable parameter λ and the smoothing parameter σ. We represent the following pseudocode to address the algorithm:

Algorithm 1 Discrete anisotropic diffusion

procedure ANISOTROPIC DIFFUSION(image, t, σ, λ)
    Transform the *image* to vector u and set current time as $t_c := 0$
    while $t_c < t$ **do**
        Calculate u_σ and then u_{σ_x}, u_{σ_y}
        Calculate the gradient orientation descriptor $J(\nabla u_\sigma)$
        Calculate $A(u)$ utilizing $J(\nabla u_\sigma)$
        Set the time-step τ such that $\tau < \frac{1}{\max\limits_{i \in J} |a_{ii}(u)|}$
        Update u with $(I + \tau A(u)) \cdot u$
        Update t_c with $t_c + \tau$
    Transform u to *image_new*
    Return *image_new*

The algorithm is construed by using the Python programming language. The necessary Python modules required are *numpy* and *scipy*. *Numpy* is great in handling numeric calculations. There are implemented algorithms for calculating the derivatives of the image or obtaining the blurred version of image. However, these have not

been used in this thesis although it could make the algorithm a bit faster and easier to implement.

As $A(u^{(k)})$ is a $nm \times nm$ matrix and very sparse, we use *scipy* module as it provides support for sparse matrices. Thus the memory of the computer does not fill up during calculations.

Python module *PIL* is used to transform the initial image to a numpy array. This can also be done using scipy module, it is just up to personal preference.

The most computationally expensive task in the algorithm was to create the matrix $A(u^{(k)})$ for every iteration. The algorithm used in this work is made to be somewhat efficient. Weickert has used semi-implicit schemes in his work which allows bigger time-steps. This however means that a very large linear system of equations has to be solved.

If one would like to build a diffusion equation algorithm for a frequent use, the iterative scheme may not be the best choice because of the limitations in the time-step size. The performance of the different schemes and algorithms is not discussed here, as this work is purely to demonstrate what the diffusion filtering is capable of.

3.3 Examples

In this section we shall investigate some examples where diffusion equation is applied. The desirable properties for images are that there should be little to no noise and the edges should be sharp. We shall see that the non-linear isotropic diffusion and the anisotropic diffusion reduce the noise levels and may even sharpen the edges. We will show few comparisons of the isotropic diffusion and the anisotropic diffusion with the same parameters and diffusivity function.

We advise the reader to zoom the images freely to inspect all the details of the images as we cannot present all the possible details here.

Images used:

Sigmoid: A simple test image

Square: Another simple test image

Passion fruit tomography: A tomography image of a passion fruit filled with salt. The bright areas represent the salt. Contrast enhancement has been applied to the original image. This image has been reconstructed from 360 tomography images, which is why the image has the little wave-like details.

Woman: A portrait of a woman. This image has small details such as eyes and hair which are important to preserve. The background is already blurred, so there are no features to preserve.

Street: An image of a street. This image has edges where the grey value changes notably. Also the antennas on the rooftops are feature to preserve.

Noitapilli: An image of a horse in a forest.

I would like to thank Areta Santos for providing the three last mentioned images.

3.3.1 Interpretation of images

Edge-enhancing property

The image below demonstrates the edge-enhancing property on a sigmoid-like function provided that the diffusivity function $g(s)$ is chosen in such a way that $b(s) = 2sg'(s) + g(s) < 0$ when $s > \lambda$ for some threshold value $\lambda > 0$.

Figure 3.1: Edge-enhancing. Left: Original image. Right: Anisotropic diffusion with $t = 1000$, $\sigma = 3$, $\lambda = 10$

We see from the figure (3.1) above that the edge is clearly much sharper although not perfect at $t = 1000$.

Convergence of the diffusion

In figure (3.2) below we see the comparison of the convergence between gaussian blur, isotropic diffusion and anisotropic diffusion. In each column the top row is the initial image.

The gaussian blur is shown in the leftmost column. It is clear that edges and details are lost when gaussian blur is applied to the image.

The middle column represents the isotropic diffusion. Right away we see that certain edges are as sharp as in the original image. Of course blurring occurs in the areas where the gradient is small. In the last two images the shell of the passion fruit is gone but the white area representing the salt is untouched. The only difference between these two images is that one small white detail at the bottom is lost.

The anisotropic diffusion is shown in the rightmost column. We see that it can preserve edges but with a cost; they become rounded. Where isotropic diffusion stops diffusing when edge is detected, anisotropic diffusion continues diffusion in the direction of the edge which is the reason for the rounded edges. It is clear that the anisotropic diffusion cannot preserve small details as well, as the isotropic diffusion for a long time.

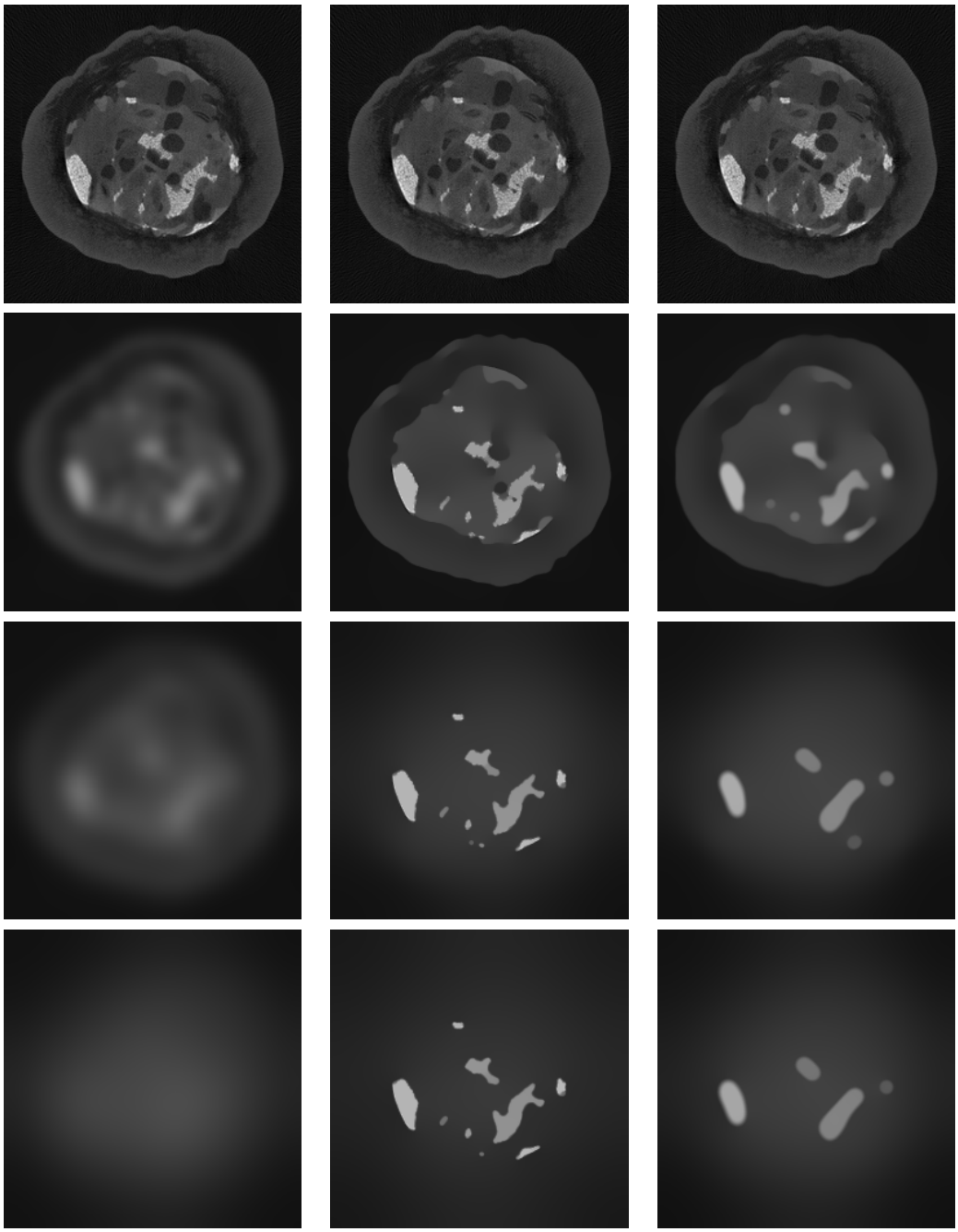

Figure 3.2: Comparison of convergences between gaussian blur, isotropic diffusion and anisotropic diffusion in this order column wise. Notice the different time scales. Gaussian blur from top to bottom: $t = 0, 20, 70, 400$. Isotropic diffusion: $\sigma = 1$, $\lambda = 4$, $t = 0, 150, 1100, 1900$. Anisotropic diffusion: $\sigma = 1$, $\lambda = 4$, $t = 0, 130, 665, 900$

Noise removal properties

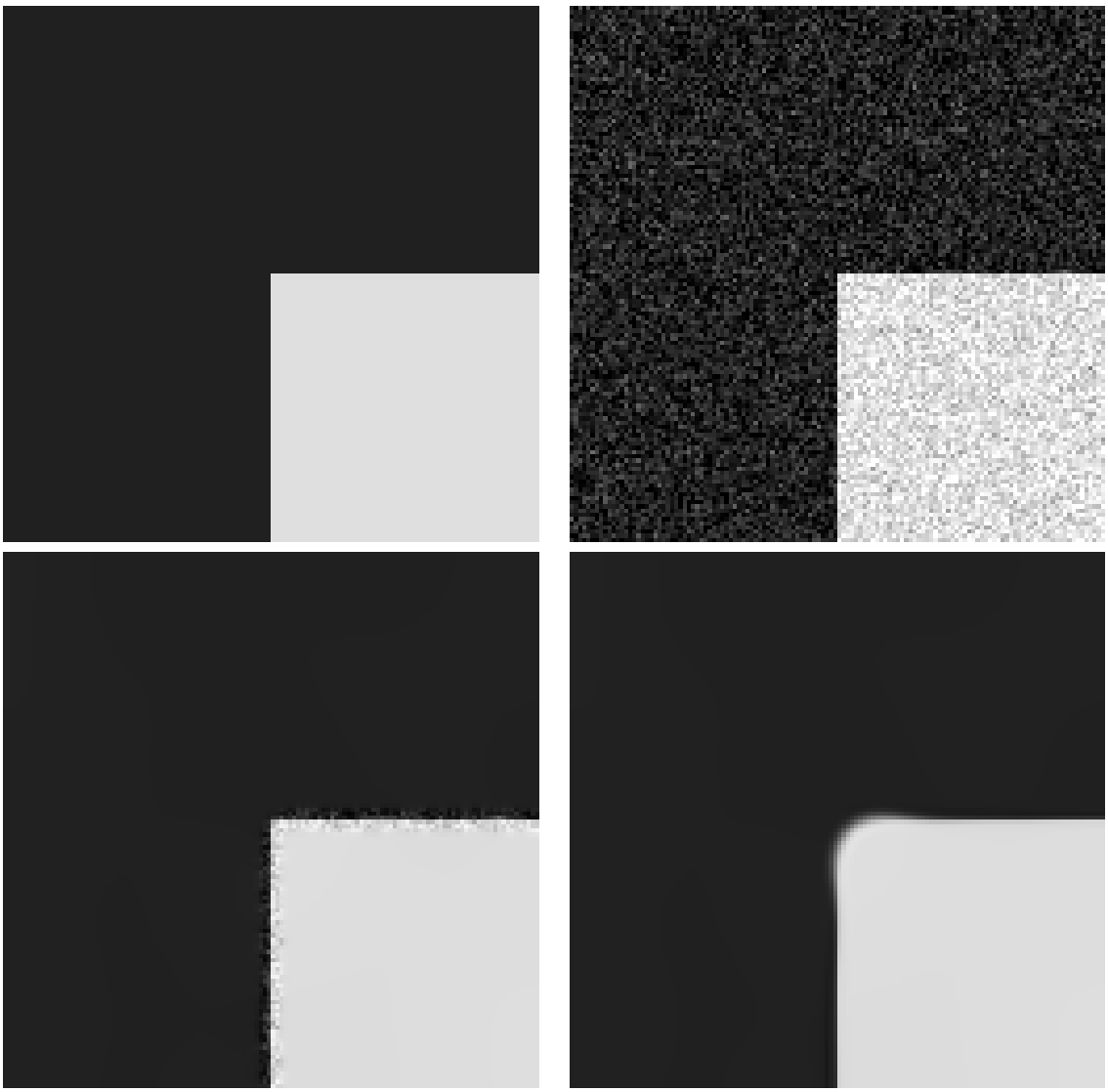

Figure 3.3: Noise removal. Top left: Original image. Top right: Original image with excessive noise. Bottom left: Isotropic diffusion $t = 100$, $\sigma = 3$, $\lambda = 8$. Bottom right: Anisotropic diffusion $t = 100$, $\sigma = 3$, $\lambda = 8$.

Figure (3.3) demonstrates the noise removal using isotropic and anisotropic diffusion on a very simple image. The bottom left image shows us that the isotropic diffusion is incapable of removing noise at the edges. However it does preserve the corner very well. The bottom right image shows that the anisotropic can deal with the noise at the edges, but the cost of this is that the corner becomes rounded. With a closer look we may see that in both cases the dark background area is not perfectly constant.

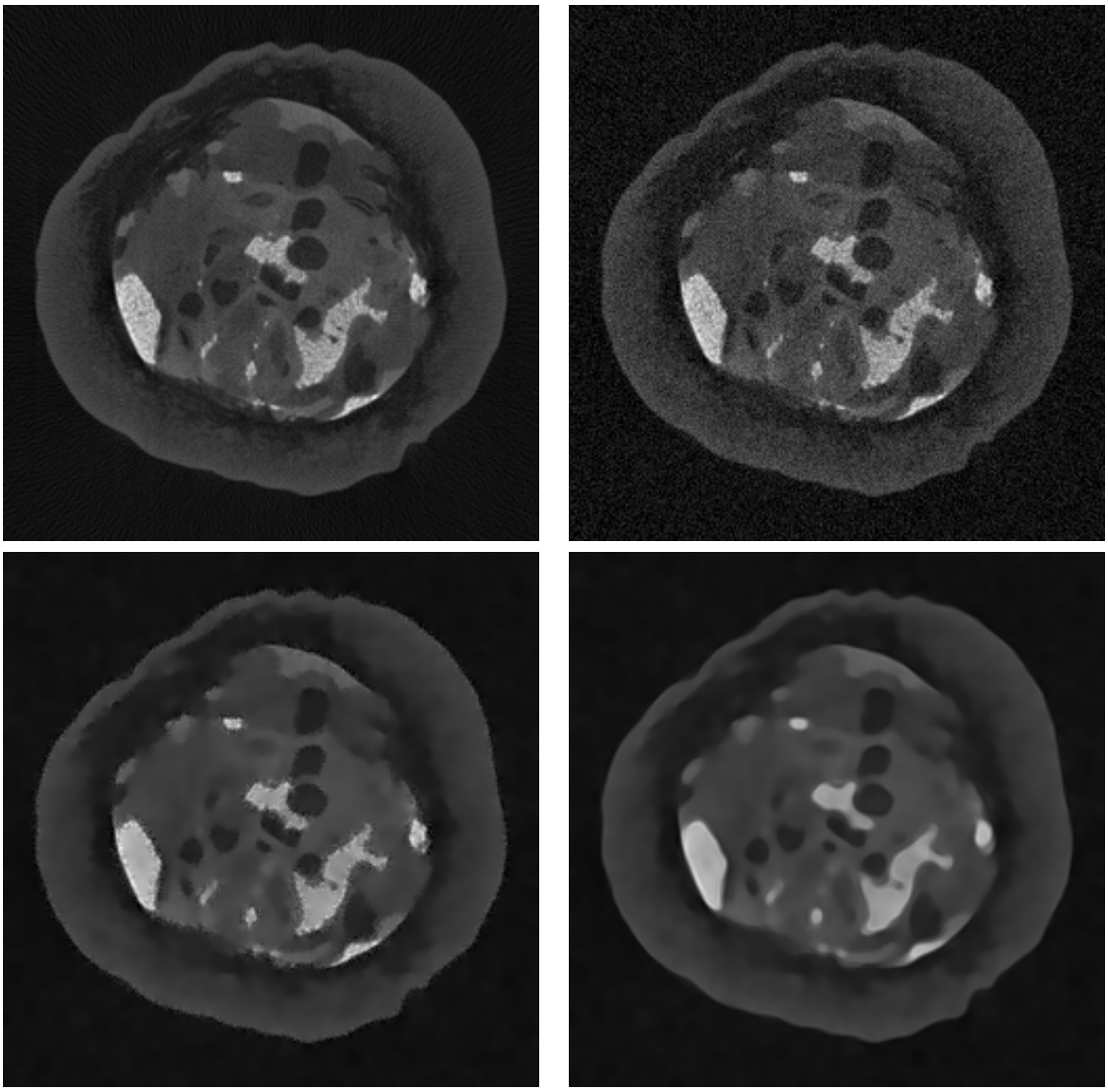

Figure 3.4: Noise removal. Top left: original tomography image. Top right: noised tomography. Bottom left: Isotropic diffusion $t = 2.6$, $\sigma = 2$, $\lambda = 5$. Bottom right: Anisotropic diffusion $t = 2.6$, $\sigma = 2$, $\lambda = 5$.

Figure (3.4) above shows isotropic and anisotropic diffusion applied to a noisy passion fruit tomography image. Again we see that the isotropic diffusion is incapable of removing the noise at the edges although it does good job in constant areas. Thus details preserve their shape, but the noise remains at the edges. The anisotropic diffusion on the other hand deals with this problem by diffusing in the direction of

edges. We see that there is little to no noise at the edges. Typical rounding of the edges occurs, but not remarkably. We see that in both cases there exist signs of random noise at the background, but this is not a big issue as there are no interesting details. It is more important what happens inside the fruit.

Figure 3.5: Noise removal of a portrait. Left: Noised image. Right: Anisotropic diffusion filtering $t = 0.3$, $\sigma = 0.8$, $\lambda = 15$. For closer look see figure (3.6) below. Parameter t was chosen that noise was reduced without losing too much details.

In the figures (3.5) above and (3.6) below are shown anisotropic diffusion applied to a photograph. In a head shot it is desired to retain every little detail possible. Background of the portrait is blurred already, so there is no need to preserve details. The amount of the noise is small but visible to naked eye, especially when zoomed. The full shot shows a reduction in the noise level while retaining the image structure. There is a very small level of blurriness but altogether the image is sharp. By looking at the hair it can be seen that the diffusion has been applied.

A closer look of the head shot (3.6) shows that the noisiness is reduced but not completely gone. This can be seen very well by looking at the shadow in the neck area. On the other hand, some of the details are disappeared. For example the hair is too blurred compared to the original image. In addition, the eye is not as sharp as in the original image but the pupil and the little reflection above it can still be distinguished. As human skin is not even, it is hard to say whether the little details

appearing in the forehead section are noise or natural differences in skin tone. The hairline in the forehead shows that the edge is not blurred.

There was no scientific method to decide at which time scale filtered image is closest to the original noiseless image. These parameters were chosen because they looked good to the naked eye. If the smoothing parameter σ was chosen too big, the pupil and the hair merged into the surroundings and they were not detected as edges. If λ was chosen too small, the noisiness did not disappear enough. By letting the time parameter t to grow bigger, the noise would be lost as would the small details. The diffusion was stopped at time $t = 0.3$ because the noise level was reduced without losing too much details.

Figure 3.6: A closer look of noise removal of the portrait. Left: Noised image. Center: Original image. Right: Anisotropic filtering $t = 0.3$, $\sigma = 0.8$, $\lambda = 15$

Anisotropic diffusion as an artistic filter

Next we show that anisotropic diffusion can be used as an image filter to make the images have a painted-like look. All the images below are at timescale $t = 5$ and $\sigma = 0.5$. For *Street* and *Noitapilli* the choice $\lambda = 20$ is made and for *Woman* we choose $\lambda = 10$.

Figure 3.7: Left: Original image. Right: Anisotropic diffusion with $t = 5$, $\sigma = 0.5$, $\lambda = 20$

Figure 3.8: Left: Original image. Right: Anisotropic diffusion with $t = 5$, $\sigma = 0.5$, $\lambda = 20$

Figure 3.9: Left: Original image. Right: Anisotropic diffusion with $t = 5$, $\sigma = 0.5$, $\lambda = 10$

From the above images we see that this kind of diffusion diminishes small details but the edges are preserved. This gives the images the painted-like look. These images also represent what would happen to the images in noise reduction problems if one would let the timescale to grow large. As there is the diffusivity function $g(s)$ and three parameters to be chosen there are numerous different results to obtain by using diffusion filtering, provided that $g(s)$ is chosen as proposed earlier.

In the original image of the *Street* the plagues were readable but not in filtered image, although it is clear that there has been text. The bricks on the dark house are no longer recognizable. We see that the rough shapes of most of the objects are preserved. Even the antennas on the rooftops are still recognizable.

Figure (3.8) shows that the flora on the image gets the painted-like look. The horse is still distinguishable with the small details of the face and the muscles still standing out.

When diffusion is applied to the portrait the face and the hair becomes much more smoothed but the mouth, the nose and the eyes preserve their shape. Pupils in the eyes are still distinguishable. We see little changes at the background, in the shadows and on the shirt.

Bibliography

[1] Joachim Weickert, *Anisotropic Diffusion in Image Processing*, B.G. Teubner Stuttgart, 1998

[2] Lawrence C. Evans, *Partial Differential Equations*, American Mathematical Society, 1997

[3] Pietro Perona, Jitendra Malik, *Scale-Space and Edge Detection Using Anisotropic Diffusion*, Transactions on pattern analysis and machine intelligence, 1990

[4] Gilles Aubert, Pierre Kornprobst, *Mathematical Problems in Image Processing*, Partial Differential Equations and the Calculus of Variations, Second Edition, Springer, 2006

[5] Francine Catté, Pierre-Louis Lions, Jean-Michel Morel, Tomeu Coll *Image Selective Smoothing and Edge Detection by Nonlinear Diffusion* Siam J. Numer. Anal. Vol 29 1992

[6] Philip Hartman, *Ordinary Differential Equations*, John Wiley & Sons, Inc, 1964

[7] Wolfgang Walter, *Ordinary Differential Equations*, Springer-Verlag New York, Inc, 1998